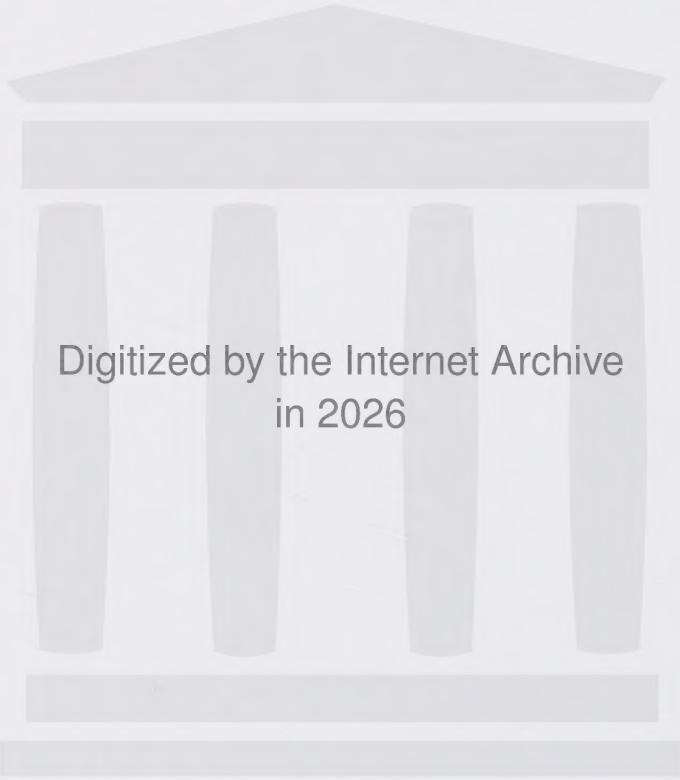


**ANSÄTZE UND VERFAHREN ZUR ANALYSE UND
NUMERISCHEN KLASSIFIKATION
PSYCHOPHYSIOLOGISCHER SIGNALE**

**Dargestellt an einem Beispieldatensatz aus der
Biofeedback-Forschung**

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der
Philosophischen Fakultät
der
Westfälischen Wilhelms-Universität
zu
Münster (Westf.)

vorgelegt von
Klaus Langer
aus Soest (Westf.)
1978



Digitized by the Internet Archive
in 2026

ANSÄTZE UND VERFAHREN ZUR ANALYSE UND NUMERISCHEN
KLASSIFIKATION PSYCHOPHYSIOLOGISCHER SIGNALE -

DARGESTELLT AN EINEM BEISPIELDATEN -
SATZ AUS DER BIOFEEDBACK-FORSCHUNG

Inaugural - Dissertation
zur Erlangung des Doktorgrades
der

P h i l o s o p h i s c h e n F a k u l t ä t

der
Westfälischen Wilhelms - Universität
zu
Münster (Westf.)

vorgelegt von

Klaus Langer
aus Soest (Westf.)

1978



Tag der mündlichen Prüfung: 14. 7. 1978

Dekan: Prof. Dr. M.E. BROCKHOFF

Referent: Prof. Dr. D. VAITL

Korreferent: Prof. Dr. H. KALLINA

VORBEMERKUNG

Daß die auf dem Hintergrund der sich abzeichnenden Formalisierungs- und Mathematisierungstendenzen der Wissenschaften, speziell der Psychologie, aufgeworfenen Fragen- und Problemkomplexe nur fragmentartig aufgegriffen, beschrieben und diskutiert werden können, ist bereits aus der auf primär methodisch-methodologische Aspekte eingeschränkten Themenformulierung zu ersehen. Diese Ausarbeitung stellt also einen Versuch dar und ist als ein Betrag zu werten, die zentrale Problematik einer systematischen Integration methodischer und substanzwissenschaftlicher Elemente nach begründeten Kriterien einer zumindest ansatzweisen Lösung zuzuführen.

Bei aller sich ergebenden Vorläufigkeit und Einschränkung bin ich mehreren Personen sehr zu Dank verpflichtet, die durch ihre motivierenden Hinweise Form und Inhalt dieser Arbeit beeinflußt haben. Neben der Unterstützung, die mir im Rahmen der Graduiertenförderung zuteil wurde, ist diese Arbeit durch die Einrichtungen und Mittel des Rechenzentrums der Universität Münster gefördert worden. So ist den Mitarbeitern des Rechenzentrums insbesondere dafür herzlich zu danken, daß die umfangreichen numerischen Auswertungen auch zu ungewöhnlicheren Zeiten durchgeführt werden konnten.

Daß die einzelnen Ansätze und Verfahren nicht isoliert, sondern in Bezug auf psychophysiologische und damit inhaltliche Punkte entwickelt und dargestellt werden konnten, verdanke ich der Unterstützung durch Herrn Prof. Dr. D. Vaitl, dessen Hinweise den Fortgang dieser Arbeit wesentlich gefördert haben. Für freundschaftliche Anregung und Kritik aus der Sicht eines Naturwissenschaftlers bin ich Herrn Dr. rer. nat. D. Steinhausen zu Dank verpflichtet, während mich Herr Dr. phil. G. Hörman auf psychologische Unstimmigkeiten und mögliche Verkürzungen aufmerksam machte.

Besonders bleibt Frau M. Geisler zu danken, die es übernahm, die schwer entzifferbaren mathematischen Textstellen in die vorliegende maschinenschriftliche Fassung zu übertragen.

GLIEDERUNG

		Seite
1	Einleitung	1
1.1	Ebenen psychophysiologischer Analyse	1
1.2	Zielfragen der Biofeedback-Forschung	2
1.3	Problemstellung	5
1.4	Aufbau der Arbeit	7
2	Signaldarstellung im Zeitbereich	11
2.1	Präzisierung der Problemstellung	11
2.2	Definition eines psychophysiologischen Signals	13
2.3	Grundlegende Begriffe und Definitionen	17
2.4	Auto- und Kreuzkorrelationsfunktion	23
2.5	Ansätze der Parameterextraktion	26
2.5.1	Einführung	26
2.5.2	Allgemeine Voraussetzungen	27
2.5.2.1	Norm	28
2.5.2.2	Orthonormalsystem	30
2.5.3	Problemunabhängige Parameterextraktion	33
2.5.3.1	Trigonometrische Funktionen	34
2.5.3.2	Walsh-Funktionen	35
2.5.3.3	Beziehung zwischen den trigonometrischen und Walsh-Funktionen	37
2.5.3.4	Andere orthogonale Funktionen	39
2.5.4	Problemabhängige Parameterextraktion	40
2.5.5	Segmentweise Parameterextraktion	42
2.5.5.1	Einführung	42
2.5.5.2	Kubische Splines	43
3	Signaldarstellung im Frequenzbereich	47
3.1	Einführung	47
3.2	Fouriertransformation	47
3.2.1	Definition	47

	Seite
3.2.2	Eigenschaften 50
3.2.3	Berechnung 52
3.3	Convolution und Spektrum 55
3.4	Spektrale Leistungsdichte und Korrelationsfunktion 57
3.4.1	Autokorrelationsfunktion und spektrale Leistungsdichte 57
3.4.2	Kreuzkorrelationsfunktion und spektrale Kreuzleistungsdichte 59
3.5	Lineare Filterung 63
3.6	Aspekte der numerischen Auswertung 66
3.6.1	Abtastung und Segmentierung 66
3.6.2	Schritte der Berechnung 67
4	Numerische Klassifikation 69
4.1	Einführung 69
4.1.1	Problemstellung 69
4.1.2	Präzisierung der Problemstellung 70
4.2	Distanzfunktionen 74
4.2.1	Definitionen 74
4.2.2	Nicht-metrische Distanzfunktionen 76
4.2.3	Metrische Distanzfunktionen 78
4.2.4	Distanzfunktionen bei Klassen 81
4.3	Verfahren 83
4.3.1	Systematisierungskriterien 83
4.3.2	Hierarchische Verfahren 84
4.3.2.1	Agglomerative Verfahren 84
4.3.2.2	Divisive Verfahren 86
4.3.3	Verfahren zur Verbesserung einer Zielfunktion 87
4.3.3.1	Einführung 87
4.3.3.2	Zielfunktionen 88
4.3.3.3	Sift-and-Shift Verfahren 91
4.3.3.3.1	Iteratives Minimaldistanzverfahren 91
4.3.3.3.2	Austauschverfahren 92
4.3.3.3.3	Minimaldistanzverfahren und Austausch- verfahren für andere Zielfunktionen 92

	Seite
4.3.4 Andere Verfahren	94
4.3.4.1 Q-Analyse	94
4.3.4.2 Konfigurationsfrequenzanalyse	96
4.3.4.3 Analyse der Punktdichte	98
 5 Exemplifikation - Analyse und Klassifikation eines Beispieldatensatzes aus der Biofeed- back-Forschung	100
5.1 Einführung	100
5.2 Aspekte der Biofeedback-Forschung	100
5.3 Skizzierung der Untersuchungs- fragestellungen	103
5.4 Schritte der numerischen Auswertung	107
5.4.1 Datenorganisation	107
5.4.2 Datendeskription	108
5.4.3 Datenklassifikation	108
5.5 Darstellung und Diskussion der Ergebnisse ..	109
5.5.1 Datendeskription	109
5.5.2 Datenklassifikation	114
 6 Weiterführende Analysestrategie	127
6.1 Einführung	127
6.2 Mathematische Aspekte	128
6.2.1 Grundbegriffe	128
6.2.1.1 Differentialgleichungen	128
6.2.1.2 Lineare Differentialgleichungen und ein Differentialgleichungssystem n-ter Ordnung ..	129
6.2.2 Stabile Zustände	132
6.2.2.1 Einführung	132
6.2.2.2 Definitionen	133
6.2.2.3 Fallunterscheidungen und graphische Darstellung der Lösungskurven	135
6.3 Konkretisation	137
6.3.1 Einführung	137
6.3.2 Darstellung und Diskussion der empirischen Daten	138

		Seite
7	Zusammenfassung	143
8	Anhang	144
8.1	Programme	144
8.1.1	Programme zur Datenorganisation	144
8.1.2	Programme zur Frequenz- bereichsdarstellung	149
8.1.3	Programme zur graphischen Ausgabe ("Hidden-Line")	154
8.1.4	Programme zur Datenklassifikation	157
8.1.5	Programme zur Visualisierung der Lösungskurven	169
8.2	Anmerkungen	173
9	Literaturverzeichnis	191

1 EINLEITUNG

1.1 EBENEN PSYCHOPHYSIOLOGISCHER ANALYSE

Allgemeiner Konsens dürfte hinsichtlich der Abgrenzung und Einordnung der dieser Ausarbeitung zuzuordnenden Forschungsrichtung bestehen. Auf dem Grenzgebiet zwischen Psychologie und Medizin hat sich als wissenschaftshistorisch vergleichsweise junge Disziplin die Psychophysiologie etabliert ^(1/1).

Orientiert an einer zum Teil vehement artikulierten Kritik über die wissenschaftlich vordergründige sowie klinisch wenig hilfreiche und aussagekräftige Dichotomisierung in ein Modell medizinisch-somatischer Krankheiten einerseits und ein Modell psychosozialer Störungen andererseits ^(1/2) werden in Überwindung perspektivisch eingeschränkter Formen wissenschaftlicher Vorgehensweise folgende vier grundlegende Analyseebenen in der Psychophysiologie unterschieden ^(1/3):

- die organisch-physiologische Ebene
- die motorisch-verhaltensmäßige Ebene
- die subjektive Erlebnisebene
- die symbolisch-kommunikative Ebene .

Im Unterschied zur rein somatischen Betrachtung dysfunktionaler Phänomene sind in der Psychophysiologie explizit die motorisch-verhaltensmäßigen, subjektiven und symbolisch-kommunikativen Momente Gegenstand der Analyse.

Psychologische Konzepte werden dahingehend erweitert, daß unter Aufhebung kategorialer Differenzen organisch-physiologische Prozesse in ihrer Funktionalität bzw. Auswirkung reflektiert und in die Konstruktion von Hypothesen und Theorien miteinbezogen werden.

Neben theoretischen Ansätzen, in denen die sich eröffnenden Möglichkeiten integrativer Analyse zur Lösung komplexer psychophysiologischer Phänomene aufgegriffen und tendenziell genutzt werden konnten ^(1/4), besitzt besondere klinische aber auch allgemeine Re-

levanz (1/5) jene theoriegeleitete Entwicklung psychophysiologischer Therapieprogramme, die oft vereinfachend unter den Begriff des "Biofeedbacks" subsumiert wird und unter dieser Bezeichnung auch eine beachtenswerte Publizität erlangen konnte (1/6).

Einen Ausgangspunkt hierzu bildeten die in Tierversuchen gewonnenen Ergebnisse zur operanten Konditionierung der vom autonomen Nervensystem regulierten Körperfunktionen, welche geringfügig modifiziert auf den Humanbereich übertragen wurden (1/7). Diesem korrespondierte, daß das bisher vorherrschende Paradigma der Unkontrollierbarkeit autonomer Funktionen und Prozesse durch willkürliche Intervention und deren reflexhafte Steuerung entscheidend in Frage gestellt wurde (1/8).

Die in speziellen Arbeiten erzielten und in der Anfangsphase eher euphemistisch beurteilten Befunde über den Erfolg lerntheoretisch ausgerichteter Therapien psychophysiologischer Symptome beziehen sich auf Veränderungen physiologischer Variablen wie Hirnstromaktivität, periphere ektodermale Reaktionen, gastritische Mobilität, Nierenfunktion, Blutdruck oder etwa Herzfrequenz (1/9).

Angesichts der sich häufenden psychischen und physischen Störungen (1/10), für die von der Allgemeinmedizin häufig nur relativ grobe Klassifizierungen und Etikettierungen wie vegetativ bedingte Störung, vegetative Dysfunktion oder z.B. vegetative Dystonie bereitgestellt werden, sind in diese therapeutischen Ansätze, die das allgemeine lerntheoretische Prinzip der Rückmeldung zur Kontrolle körpereigener Funktionen zu nutzen suchen, nicht geringe Erwartungen und Hoffnungen gesetzt worden.

1.2 ZIELFRAGEN DER BIOFEEDBACKFORSCHUNG

Im Unterschied zu der eingangs konstatierten Übereinstimmung hinsichtlich der Einordnung der Psychophysiologie in eine Wissenschaftssystematik existieren über die klinisch-psychologische Bedeutung psychophysiologischer Therapiemodelle stark kontroverse und divergierende Ansichten, die von einer Einschätzung als zukunftsweisender, erfolgreicher Intervention bis hin zum Vorwurf

reichen, lediglich bereits Bekanntes zu adaptieren, ohne hierbei ein schlüssiges Strukturierungsprinzip angeben und begründen zu können.

Insofern es sich bei der Biofeedback-Therapie um eine relativ junge Forschungsrichtung handelt, dürfte eine abschließende und die eigentliche kritische Auseinandersetzung vermeidende Beurteilung sicherlich verfehlt sein, wenn primär darauf verwiesen wird, daß nur minimale Änderungen in den betrachteten physiologischen Größen erreicht worden seien.

Angemessener scheint demgegenüber die ernüchternde Skepsis zu sein, in der auf die in speziellen Versuchsserien erhobenen Befunde aufmerksam gemacht wird und in deren Anschluß sich präzisierbare Fragen ergeben. Ist man an einer derartigen in praktischer Absicht formulierten Kritik psychophysiologischer Untersuchungen interessiert, führt dies auf eine Differenzierung in drei Zielebenen, die gleichsam vertikal zu den skizzierten Ebenen psychophysiologischer Analyse anzuordnen sind ^(1/11):

- "In welchem Ausmaß kann die Rückmeldung zu einer Veränderung des untersuchten biologischen Prozesses führen (Effekt-Nachweis)?
- Welche psychologischen Variablen lassen sich isolieren, die unter Rückmeldebedingungen eine Funktionsänderung dieses biologischen Prozesses hervorrufen (Effekt-Determination)?
- Inwieweit kann dieses Verfahren klinisch zur Behandlung von psychophysiologischen Dysfunktionen eingesetzt werden (Effekt-Generalisation)?"

Entscheidend für eine differenzierte Beantwortung der auf diesen Zielebenen psychophysiologischer Forschung aufgeworfenen Fragenkomplexe ist zunächst, daß die Bedingtheit der in den Untersuchungen zum Biofeedback erzielten Resultate durch die implizit vorhandenen Prämissen der formalen Konstruktion reflektiert und in ihrer Auswirkung auf die inhaltliche Interpretation in Betracht gezogen wird. Erst wenn methodische und methodologische Probleme über das Vorgehen oder den Weg, zu bestimmten Sätzen und Aussagen

zu gelangen, in Angriff genommen werden, wird man die registrierte mangelnde Aussagekraft vorliegender Arbeiten zum Biofeedback überwinden und auf die Fragen des Effekt-Nachweises, der Effekt-Determination sowie der Effekt-Generalisation eine wissenschaftlich begründete Antwort finden können.

Neben methodologischen Erörterungen, die Prinzipien einer adäquaten experimentellen Instrumentation einschließen, nehmen daher methodische Aspekte über geeignete mathematische und statistische Verfahren und Modelle zur Beschreibung und Analyse psychophysiologischer Funktionsabläufe eine besondere Stellung ein.

Abhängig von den Spezifika des untersuchten Gegenstandsbereichs, die mit den Mitteln klassischer Statistik nur höchst unzureichend erfasst werden können, erlangen alternative und in der Psychologie bislang kaum verbreitete methodische Ansätze steigende Bedeutung, welche die für psychophysiologische Variablen kennzeichnenden Unterschiede hinsichtlich der zeitlichen Latenzen berücksichtigen (1/12).

So nehmen etwa Variablen der Gehirn-, Herz- oder Muskeltätigkeit innerhalb von Millisekunden neue Werte an, während peripher physiologische Merkmale sich nach Sekunden oder Minuten verändern. Die Registrierung von Änderungen biochemischer Größen erfordert Meßintervalle im Minuten-, nicht selten jedoch Stundenbereich. Ausgesprochen unterschiedlich und zum Teil schwierig ist die angemessene Erfassung von Daten auf der Verhaltens- und Erlebnisebene, da je nach definierter Untersuchungseinheit sehr verschiedene Methoden und damit Meßzeiten zugrundegelegt werden (1/13).

Aufgrund der Komplexität des Untersuchungsgegenstandes, die bereits durch die obige Strukturierung in die einzelnen psychophysiologischen Analyseebenen sichtbar geworden sein mag, führt zudem erst die simultane Aufnahme mehrerer Meßgrößen zu relevanten Aussagen. Charakteristisch für die Psychophysiologie ist, daß physiologische Funktionen in Relation zu psychologischen Variablen meist erst in Form von spezifizierten Variablenkonfigurationen oder Mustern zu identifizieren sind, wie dies paradigmatisch an

den Kategorien 'Angst', 'Streß' oder 'Motivation' deutlich wird (1/14). Die Erfassung dieser psychophysiologischen Reaktions- und Verhaltensmuster und deren angemessene Interpretation setzt schließlich den gezielten und methodisch begründeten Einsatz multivariater Verfahren und Modelle voraus (1/15).

1.3 PROBLEMSTELLUNG

Aus der hier notwendig fragmentarischen Darstellung der Ebenen psychophysiologischer Analyse und der sich stellenden Zielfragen ist die dieser Ausarbeitung zugrundeliegende Problemstellung abzuleiten, insofern ausgehend von den noch offenen Fragen und den sich abzeichnenden methodischen und methodologischen Problemen und Defiziten (1/16) ein Versuch zu einer genaueren und an einem speziellen psychophysiologischen Beispieldatensatz konkretisierten Diskussion differenzierterer methodischer Aspekte unternommen wird.

Zur Veranschaulichung der im folgendem aufgegriffenen Problematik, die einen relativ exakt definierbaren Bereich innerhalb des psychophysiologischen Forschungsgebiets umfaßt (1/17), werde eine physiologische Größe oder Variable betrachtet, an welche ursprünglich die Hoffnung geknüpft wurde, sie als Abbild psychischer Prozesse ansehen zu können (1/18).

In der äußeren Rindenschicht des Neocortex entstehen durch synchronisierte neuronale Aktivitäten elektrische Potentialdifferenzen, die mit Hilfe geeigneter Verstärker und Registriergeräte aufgenommen werden können. Diese bioelektrische Hirnaktivität, die in Form des Elektroencephalogramms (EEGs) aufgezeichnet wird, stellt ein wichtiges diagnostisches Hilfsmittel traditionellerweise in der Medizin, insbesondere der Neurologie, Neurochirurgie und Pädiatrie, aber in zunehmendem Maße auch in der Psychophysiologie dar. Wichtigstes diagnostisches Kriterium ist in der Regel die Frequenz und man unterscheidet vier EEG-Frequenzbereiche oder Frequenzbänder (1/19).

Eine abgesicherte Interpretation und Beurteilung des EEGs anhand der aufgezeichneten Originalkurven ist jedoch ausgesprochen schwie-

rig und hängt stark von subjektiven Aspekten ab, was schon durch die hohe Messwertanzahl bedingt ist ^(1/20). Oft wird durch eine Aufzeichnung der Daten im Zeitbereich ^(1/21) zudem das optische Auflösungsvermögen deutlich überschritten.

Die sich daraus ergebende Notwendigkeit, effektive Darstellungsformen zu finden, führt zu einer Transformation der Daten in den Frequenzbereich. Mit Hilfe spezieller Verfahren und Methoden werden die in den Ausgangsdaten enthaltenen Anteile etwa hinsichtlich der Frequenz, Amplitude oder Phase ausgewertet. Auf dieser Grundlage ist unter gewissen Bedingungen bereits eine vergleichsweise reliable visuelle Beurteilung des EEGs möglich.

Dieser Übergang von der Zeit- zur Frequenzbereichsdarstellung impliziert eine wesentliche Reduzierung des zu betrachtenden Datenmaterials. Während insbesondere beim EEG die Spektralbeschreibung sinnvolle Rückschlüsse erlaubt, kann man eine entsprechende Datenreduktion bei anderen physiologischen Variablen auch mit Hilfe alternativer Verfahren und Modelle erreichen, die allgemein unter der Bezeichnung "Parameterextraktionsverfahren" subsumierbar sind ^(1/22).

Die durch Parameterextraktion oder Übergang zur Frequenzbereichsanalyse mögliche Verringerung der Informationsmenge ist allerdings in vielen Fällen noch nicht ausreichend und man wird weitere Parameter "höherer Ordnung" suchen, deren Anzahl gering ist und die zugleich eine differenzierte, klinisch relevante Dateninterpretation gestatten.

Für das EEG berechnet man z.B. verschiedene Spektralparameter in den vier standardisierten Frequenzbändern. Ermittelt wird pro Band üblicherweise die dominante Frequenz, die absolute und relative Ausprägung, die Percentilgrenzen, die äquivalente Bandbreite und der sogenannte Ausprägungsquotient ^(1/23).

Die so gewonnenen Parameter bilden dann die Basis für die eigentliche statistische Analyse. Zur quantitativen Datenauswertung bieten sich neben bekannten multivariaten diskriminanzanalytischen und varianzanalytischen Ansätzen insbesondere Verfahren der numerischen Klassifikation an, insofern diese unmittelbarer auf die

inhaltliche Problemstellung ausgerichtet sind, eine adäquate Strukturierung der Daten zu konstruieren, die auf unterscheidbare Klassen im Sinne spezifischer Symptom-, Diagnose- und Therapiecluster führt.

Bei der EEG-Analyse wird man beispielsweise ausgehend von den Spektralparametern möglichst homogene Teilmengen intra- und interindividueller Funktionsverläufe identifizieren wollen und die dort vorzufindenden relativ konsistenten Verlaufsmuster zu existierenden inhaltlichen Hypothesen und Theorien in Beziehung setzen.

Dieser Strategie, von der Zeitbereichs- zur Frequenzbereichsdarstellung überzugehen und darauf aufbauend geeignete Parameter für eine anschließende multivariate Auswertung zu extrahieren, entspricht formal die schrittweise Verdichtung der Information. Es handelt sich hier also um eine sukzessive Datenreduktion, die nicht nur für das EEG, sondern mit gewissen Modifikationen ebenso erfolgsversprechend für andere psychophysiologische Signalfunktionen zu nutzen ist.

In der Abb. 1.3-1 werden die bisher skizzierten Schritte wiedergegeben. Die in der Ordinate aufzutragende Informationsmenge (etwa in bit) ist hier nicht spezifiziert worden, da je nach involvierter Variable die Angaben stark schwanken. In der Literatur werden Faktoren von 10^{-5} - 10^{-8} angegeben (1/24).

Generell bleibt zu betonen, daß in der Regel eine Analyse der Signalfunktion im Frequenzbereich erst in Verbindung mit weiterführenden und anspruchsvolleren mathematisch-statistischen Auswertungsverfahren sinnvoll ist und der Untersuchungsfragestellung angemessen sein dürfte.

1.4 AUFBAU DER ARBEIT

Aus den in der Abb. 1.3-1 skizzierten einzelnen Phasen der schrittweisen Informationsverdichtung ergibt sich die in dieser Arbeit angezeigte Gliederung des Textes.

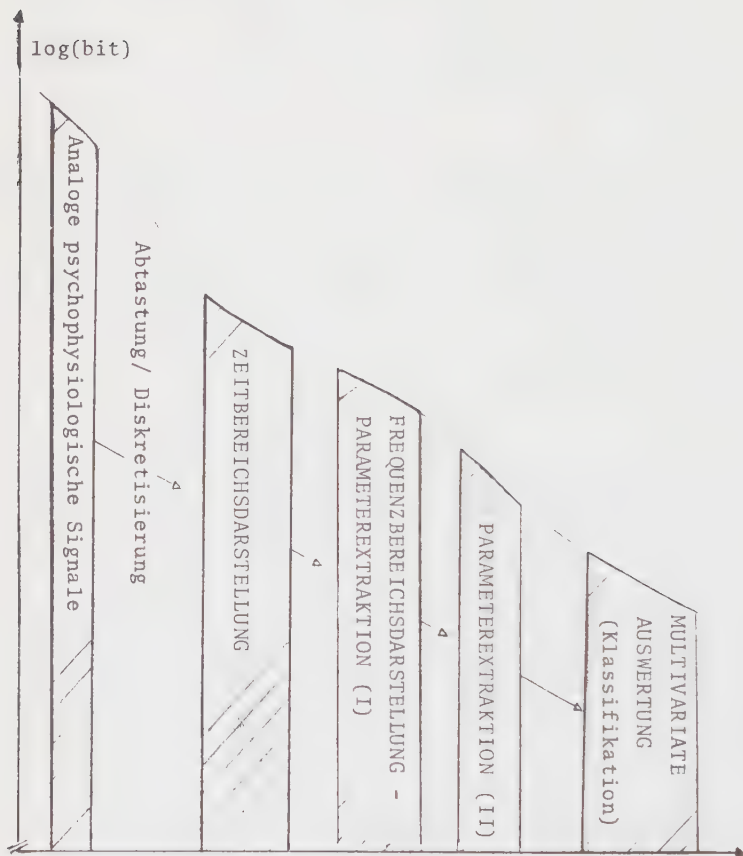


Abb. 1.3-1

SUKZESSIVE INFORMATIONSREDUKTION

(Analyse im Zeit- und Frequenzbereich und anschließende multivariate Auswertung)

Nach der Präzisierung der Problemstellung im Punkt (2.1) werden zunächst Ansätze und Verfahren zur Signaldarstellung im Zeitbereich diskutiert. Einen größeren Raum nehmen dabei Fragen einer adäquaten Parameterextraktion ein, auf die unter Bezug auf mehr mathematische Aspekte eingegangen wird. Dieser Zugang wird aus Gründen der Transparenz und Lesbarkeit des Textes bevorzugt. So wird grundsätzlich auch im weiteren Verlauf eine einheitliche Notation und ein Kompromiß zwischen formaler Exaktheit und - bezogen auf den Adressatenkreis ^(1/25) - Verständlichkeit der Darstellung angestrebt.

In dem anschließenden Punkt über Methoden zur Signaldarstellung im Frequenzbereich finden sich die für psychophysiologische Untersuchungen und Projekte relevanten Definitionen und Begriffe. Nach ausführlicherer Behandlung der Fouriertransformation werden die wichtigsten digitalen Methoden zur spektralen Signalanalyse aufgegriffen und erläutert.

Der vierte Punkt enthält eine verkürzte Erörterung jener Verfahren der numerischen Klassifikation, die für die Praxis geeignet und daher in dem hier diskutierten Kontext der Signalanalyse wachsendes Interesse besitzen dürften.

Um den Konnex zur Psychophysiologie zu betonen, werden in dem fünften Punkt ausgewählte Methoden der Signalanalyse und -klassifikation auf einen Beispielsdatensatz zur Biofeedback-Forschung angewandt und im Hinblick auf ihre Leistungsfähigkeit und Aussagerelevanz diskutiert.

Mit dem Ziel, in einer Art Ausblick oder Entwurf eine beispielhafte Strategie für mögliche Modellkonstruktionen zur Erfassung komplexer psychophysiologischer Phänomene anzugeben, wird unter Punkt 6 mit Hilfe von Methoden der interaktiven grafischen Datenverarbeitung eine mit vertretbarem Rechenaufwand realisierbare Strategie zur Visualisierung stabiler Zustände entwickelt und auf die unter Punkt 5 beschriebenen Beispieldaten angewandt.

Aus didaktischen Überlegungen wird im laufenden Text auf die Angabe der benutzten Programme zur Datenorganisation und -auswertung

verzichtet. Diese werden im Anschluß an die Zusammenfassung unter Punkt 7 im achten Gliederungspunkt nach einheitlichen Standards verfaßt aufgelistet. Desweiteren finden sich die Anmerkungen zu den einzelnen Punkten im Anhang. Das Literaturverzeichnis bildet den formalen Abschluß dieser Arbeit.

2 SIGNALDARSTELLUNG IM ZEITBEREICH

2.1 PRÄZISIERUNG DER PROBLEMSTELLUNG

Von den zahlreichen und konkurrierenden Möglichkeiten der Signalanalyse und Signalklassifikation werden basierend auf der in (1.3) umrissenen Problemstellung in den folgenden Abschnitten jene Methoden, Verfahren und Modelle aufgegriffen und konkreter beschrieben, die gemäß den einleitend skizzierten Zielfragen der Biofeedback-Forschung ^(2/1) zur Analyse und Klassifikation psychophysiologischer Signalfunktionen auf verschiedenen Betrachtungs- oder Zerlegungsebenen ^(2/2) besonders angezeigt sein dürften ^(2/3).

Beschreibung und Diskussion dieser Ansätze und Verfahren zur Analyse und numerischen Klassifikation psychophysiologischer Signale können aus der in Abb. 1.3-1 wiedergegebenen Differenzierung, welche die Grundprinzipien einer gegenstandsangemessenen Datenreduktion erkennbar werden läßt, abgeleitet und strukturiert werden. So wird zunächst in Punkt 2 die Signaldarstellung im Zeitbereich behandelt. Punkt 3 umfaßt die Signaldarstellung im Frequenzbereich, während in Punkt 4 die numerische Klassifikation Gegenstand der Betrachtung und Erörterung ist.

Dieser Gliederung korrespondiert der in experimentellen Studien indizierten und daher auch in zunehmenden Maße realisierten Vorgehensweise, den Prozeß der Datenauswertung in drei allerdings nicht in jedem Fall streng voneinander zu unterscheidende Schritte oder Phasen aufzuteilen.

In einem ersten Schritt wird man die erhobenen, diskretisierten und korrigierten Meßwerte zweckmäßigerweise mit Hilfe vorwiegend deskriptiver Methoden im Zeitbereich darstellen. Aus diesen Befunden ist zumeist mit hinreichender Sicherheit zu entscheiden, welche mathematisch-statistischen Verfahren herangezogen und aufgrund welcher Kriterien diese als adäquat ausgewiesen werden können. Im zweiten sowohl forschungslogisch als auch zeitlich nachgeordneten Schritt werden komplexere mathematische und statistische Methoden der Datenauswertung eingesetzt, die unter gegebenen Be-

dingungen die Möglichkeiten einer Frequenzbereichsdarstellung nutzen. Aus dem breiten Spektrum verfügbarer Verfahren wird man in naheliegender Weise diejenigen auswählen, die in einem methodisch-systematischen Bezug zu weiterführenden multivariaten Auswertungsstrategien stehen.

Die mit diesen in der dritten Phase des Datenanalyseprozesses erzielten Aussagen stellen in der Regel die gesuchte und eigentlich forschungsrelevante Information dar und ermöglichen interessante Folgerungen. Da an dieser Stelle im Vergleich zu anderen multivariaten Verfahren wie etwa der Diskriminanz- oder Faktorenanalyse Verfahren der numerischen Klassifikation eine nicht unwichtige und in der Praxis bislang nur begrenzte Funktion übernehmen, wird die Gruppe dieser speziellen Verfahren näher behandelt.

Die vorfindbare Trennung in die drei Hauptpunkte (Signaldarstellung im Zeitbereich, Signaldarstellung im Frequenzbereich, Numerische Klassifikation) ist also inhaltlich und formal zu begründen und impliziert die jeweils weitere Differenzierung in die betreffenden Punkte und Unterpunkte. Von dort ist darüber hinaus abzuleiten, daß die eingangs formulierte und aus verschiedener Perspektive zu erörternde Problemstellung auf primär methodische und mathematisch-statistische Aspekte eingeschränkt wird.

So werden nach der Definition eines psychophysiologischen Signals die in diesem Kontext grundlegenden Begriffe erläutert. Anschließend wird auf die Auto- und Kreuzkorrelationsfunktion eingegangen. Einen breiteren Raum nehmen Methoden der Parameterextraktion ein. Diese führen zur Signaldarstellung im Frequenzbereich über, welche nach eingehender Beschreibung der Fouriertransformation die Definition zentraler Begriffe wie Convolution und Spektrum einschließt. Nachdem die Prinzipien linearer Filterung erwähnt worden sind, werden Aspekte der numerischen Auswertung behandelt. In dem anschließenden Abschnitt finden sich die wichtigsten Ansätze und Verfahren der numerischen Klassifikation. Ihre Beschreibung und Diskussion erfolgt in mehreren Schritten.

2.2 DEFINITION EINES PSYCHOPHYSIOLOGISCHEN SIGNALS

Unter einem "psychophysiologischen Signal" werde im folgenden der enger umschriebene Bereich von psychologischen, physiologischen oder biologischen Merkmalen verstanden, die als Funktion eines Parameters t darstellbar und weitgehend analysierbar sind. Wenn nicht anders angegeben, symbolisiert t die Zeit, obwohl auch andere, meist physikalische Bedeutungen zutreffen können.

Eine besondere und im Vergleich zur Allgemeinen Psychologie für die Psychophysiologie charakteristische Klasse bilden die verkürzt auch als "Biosignale" bezeichneten Variablen. Neben respiratorischen, kardiovaskulären oder motorischen Größen werden unter diesen Terminus generell biologische Signale subsumiert, welche mittels geeigneter Meßfühler registriert und durch Technik verstärkt gemessen werden können ^(2/4). In der Regel variieren diese Signale hinsichtlich Amplitude, Phase oder Frequenz in kennzeichnenden Grenzen ^(2/5). Je nach Fragestellung wird es sich um ein spezielles oder um eine simultane Aufnahme mehrerer für die Untersuchung relevanter Signale handeln.

Lediglich zur Illustration sind in der Abb. 2.2-1 wichtige psychophysiologische Signale wiedergegeben, die z.B. hinsichtlich ihrer Variabilität, Periodizität, Frequenzverteilung, Amplitudendichte oder anderen charakteristischen Kenngrößen ausgewertet werden.

Die vorliegenden Daten stellen eine Simultanaufzeichnung der Gehirn-, Herz-, Augen-, Atmungs- und Muskeltätigkeit dar ^(2/6). Zeile 0 enthält die Zeitschreibung (1 Sek.) für das ausschnitthaft angeführte Beobachtungsintervall. Zeile 1 gibt die Kurve der Herzfrequenzintegration wieder. Statt dieser Darstellung wird im klinischen Bereich zumeist die Originalschreibung des EKGs herangezogen, während man im Rahmen der Biofeedback-Forschung und ergonomischen Untersuchungen in der Regel die Intervalle zwischen den R-Spitzen betrachtet. Die Registrierung der Atem-

tätigkeit, die mittels eines Thermistors erfolgte, findet sich in Zeile 2. In Zeile 3 und 4 sind die Kurven für das EKG/Aktogramm bezogen auf die untere und obere Extremität wiedergegeben, wobei die zugehörigen Ableitungselektroden einmal an der Vorderseite der Oberschenkel und zum anderen an der Außenseite der Oberarme fixiert wurden. Zeile 5 stellt die Aufzeichnung des Kinn-Elektromyogramms dar. Dem Elektro-Oculo-Gramm (EOG) in der sechsten Zeile schließt sich in den folgenden Zeilen die 12-kanalige Registrierung der Hirntätigkeit an. Die Kleinbuchstaben (r,l) bezeichnen die jeweilige Hemisphäre, während die Abkürzung in Großbuchstaben die Elektrodenverbindungen angeben (F-T = fronto-temporal; T-C = temporal-central; C-O = central-occipital). In etwa der Mitte der EEG-Linien zwischen der Zeile 9 und 10 ist als Bezugsgröße für die Amplituden eine 100 mikro Volt Eichung eingetragen.

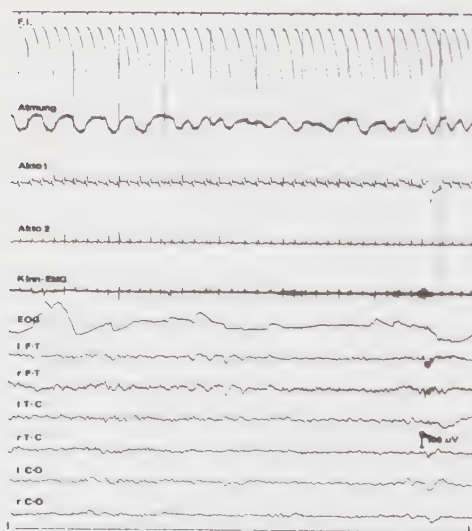


Abb. 2.2-1

SIMULTANAUFZEICHNUNG PSYCHOPHYSIOLOGISCH RELEVANTER SIGNALE

In der Regel ist der relevante psycho-physiologische Prozeß nicht direkt zugänglich, sondern wird über eine geeignet abgeleitete und verstärkte Folge von Signalen erfaßt. Dieser Sachverhalt ist neben dem in (1.3) kurz beschriebenen EEG besonders deutlich anhand des Elektrokardiogramms zu erkennen.

Der Erregungsvorgang des Herzens wird durch vegetativ-neurale und ionale Mechanismen gesteuert. Erregungs- und Kontraktionsverlauf des Herzens werden über das Prinzip der elektromechanischen Koppelung miteinander verbunden. Das EKG als Ausdruck der Erregung (und nicht der Kontraktion) kommt durch in Längsrichtung zu den Myokarfasern verlaufende Potentialdifferenzen zustande, die vom Herzen ausgehend im leitenden Medium des Körpers zwischen definierten Stellen der Körperoberfläche abgeleitet, verstärkt und aufgezeichnet werden. Man erhält also eine zeitliche Folge von Werten, die für jeden Herzzyklus ein kennzeichnendes Muster bilden.

Die spezielle Form der EKG-Kurve, die nicht nur von der Erregung, sondern auch von der Art der Ableitung abhängig ist, liefert diagnostisch verwertbare Anhaltspunkte über Frequenz, Ursache, Ausbreitung und Rückbildung der Erregung. Es dürfte jedoch deutlich sein, daß gleichartige Veränderungen im EKG verschiedene Ursachen haben können und nicht jede Störung der Herzerregung im EKG manifest wird (2/7).

Da über den physiologischen Prozeß nicht unmittelbar Aussagen gemacht werden können und speziell in der Psychophysiologie Ansätze der Systemidentifikation und Systemanalyse (2/8) vor erheblichen Schwierigkeiten stehen und zumeist heuristisch sowie unter Bezug auf eine Reihe stark vereinfachender Annahmen zu bearbeiten sind, werden an dieser Stelle zunächst einschränkend Fragen und Aspekte der Signalanalyse aufgegriffen und diskutiert.

Im Rahmen der Signalanalyse scheint es sinnvoll und zweckmäßig, die nachfolgend aufgeführten Unterscheidungen hinsichtlich der zu untersuchenden Signalfunktion zugrunde zu legen.

$$(2.2-1) \quad s(t_1) \in \{s \mid s_a \leq s \leq s_b\}$$

$$(2.2-2) \quad s(t_1) \in \{s_1, s_2, \dots, s_m\}$$

$$(2.2-3) \quad s(t_1) \in \{^q s(t) \mid q=1, 2, \dots\}$$

In (2.2-1) entstammt $s(t_1)$ einem festen Intervall. s_a, s_b stellt in diesem Fall eine kontinuierliche Zufallsvariable dar ^(2/9).

Nimmt $s(t_1)$ hingegen nur einen bestimmten Wert aus einem vorgegebenen endlichen oder abzählbaren Wertevorrat an, handelt es sich um eine diskrete Zufallsvariable.

Die durch (2.2-1) und (2.2-2) festgelegten Signalfunktionen bestehen nur aus einem Funktionswert bzw. einer Funktion. Für solche pulsformigen Signale ist folglich nur ein begrenztes Beobachtungsintervall relevant. Signale dieser Art besitzen demnach eine endliche Energie

$$(2.2-4) \quad \int_{-\infty}^{\infty} |s(t)|^2 dt < \infty$$

und eine verschwindende mittlere Leistung (s.u.).

Stellt man sich derartige Signale gleichsam nebeneinander geschaltet vor und nimmt an, daß zwischen den Einzelsignalen keine festen Beziehungen existieren, ergibt sich eine stochastische Zeitfunktion. Wird desweiteren unterstellt, daß diese Funktion als Element einer bestimmten Menge von Signalfunktionen aufgefaßt werden kann, führt dies zur Unterscheidung (2.2-3).

In Abgrenzung zu (2.2-1) und (2.2-2) wird $s(t_1)$ in (2.2-3) als Folge oder Familie von Zufallsvariablen definiert. Die spezielle, als Repräsentant oder Musterfunktion bezeichnete Funktion wird durch die Hochzahl q gekennzeichnet ^(2/10).

Eine solche Familie von Zufallsvariablen heißt vereinbarungsgemäß stochastischer Prozeß und man schreibt kurz

$$(2.2-5) \quad \gamma(t) = \{^q s(t) \mid q=1, 2, \dots; t \in T\}.$$

Ein stochastischer Prozeß wird m.a.W. durch die Gesamtheit seiner Repräsentanten gegeben.

Da sich der Wert von (t) bzw. $s(t)$ fortlaufend ändert, wird man im Unterschied zu den pulsförmigen Signalen ein unendliches Intervall zugrundelegen. Im konkreten Fall wird sich das Beobachtungsintervall auf einen längeren, meist vom Untersuchungsziel bestimmten Zeitraum erstrecken. Die mittlere Energie (2.2-4) dieser Signale ist unendlich und für die mittlere Leistung gilt

$$(2.2-6) \quad |s(t)|^2 := \lim_{P \rightarrow \infty} \frac{1}{2P} \int_{-P}^{+P} |s(t)|^2 dt$$

2.3 GRUNDLEGENDE DEFINITIONEN UND BEGRIFFE

Die zur Beschreibung von Signalfunktionen elementaren Begriffe werden zunächst anhand eines einfachen und als Idealisierung zahlreicher empirischer Phänomene benutzten Funktionstyps der Sinus- bzw. Cosinusfunktion erläutert. Nach der Definition von Beschreibungsgrößen für beliebige periodische Funktionen werden unter Bezug auf die Unterscheidungen (2.2-1/3) die für einen stochastischen Prozeß charakteristischen Größen kurz skizziert.

Ausgangspunkt der Betrachtung sei zunächst eine periodische Funktion

$$(2.3-1) \quad s(t) = s(t+P) \quad (\text{für alle } t \in \mathbb{R})$$

sehr einfacher Gestalt. Eine solche Eigenschaft besitzt etwa die Funktion

$$(2.3-2) \quad s(t) = a \sin(\omega t + \psi).$$

Die Konstante a gibt die maximale Auslenkung aus der Ruhelage an und heißt Amplitude oder Scheitelwert. Nicht selten wird a normiert, so daß

$$(2.3-3) \quad -1 \leq s(t) \leq +1 \quad \text{ist.}$$

In (2.3-1) gibt P die Periode an. Eine halbe Periode ist für (2.3-2) gerade die Zeit zwischen zwei Nulldurchgängen.

Das Inverse von P

$$(2.3-4) \quad f = 1/P$$

heißt Frequenz und wird in Hertz ($\text{Hz} = \text{sec}^{-1}$) bzw. einem Vielfachen dieser Einheit gemessen.

Berücksichtigt man den Zusammenhang zwischen einer gleichförmigen Kreisbewegung und der Sinusfunktion, entspricht die Zeit P für das Durchlaufen einer vollen Sinuslinie einem Winkel von 2π oder 360° .

Das Argument

$$(2.3-5) \quad w = 2\pi/P = 2\pi f$$

wird Kreisfrequenz oder auch Winkelgeschwindigkeit genannt. Der von der Nullstellung ($t_0=0$) in der Zeit $t = t_1 - t_0$ überstrichene Drehwinkel

$$(2.3-6) \quad \psi = 2\pi/P \cdot t = 2\pi f \cdot t = w \cdot t$$

wird als Phase der jeweiligen Momentanwerte t bestimmt. ψ hängt von der Wahl des Zeitnullpunktes ab und wird bei praktischen Berechnungen oft vernachlässigt. Damit wird (2.3-2) zu

$$s(t) = a \sin(w \cdot t) \quad \text{vereinfacht.}$$

Bei streng periodischen Funktionen gilt (2.3-1) bzw.

$$(2.3-7) \quad s(t-nP) = s(t-nP+nP) = s(t) \quad (n \in \mathbb{N})$$

mit festem P.

Für fast-periodische Funktionen

$$(2.3-8) \quad s(t+P) \approx s(t)$$

kann P sich mit der Zeit ändern.

Liegt eine beliebige periodische Funktion vor, wird die Schwankung oder Schwankungsbreite s_b als Differenz von größtem und kleinstem Wert definiert

$$(2.3-9) \quad s_b := s_{\max} - s_{\min}.$$

Der zeitliche, lineare arithmetische Mittelwert

$$(2.3-10) \quad \overline{s(t)} = \lim_{P \rightarrow \infty} \frac{1}{2P} \int_{-P}^P s(t) dt$$

heißt auch Gleichwert, Gleichanteil oder Gleichkomponente. Mit (2.3-10) läßt sich eine beliebige periodische Funktion $s(t)$ in einen sog. Gleichanteil $\overline{s}$ und einen Wechselanteil $\hat{s}$ zerlegen

$$(2.3-11) \quad s(t) = \overline{s(t)} + \hat{s}(t).$$

Aus (2.3-10) und (2.3-11) ergibt sich, daß der zeitliche lineare Mittelwert des Wechselanteils verschwindet

$$(2.3-12) \quad \lim_{P \rightarrow \infty} \frac{1}{2P} \int_{-P}^P \hat{s}(t) dt = 0.$$

Die Extremwerte des Wechselanteils bilden die Differenzen

$$(2.3-13) \quad (s_{\max} - \overline{s}) \quad \text{bzw.} \quad (s_{\min} - \overline{s}),$$

wobei die betraglich größte Differenz der beiden Werte als Scheitелwert bezeichnet wird.

Es wurde bereits erwähnt, daß der Ausdruck

$$(2.3-14) \quad \overline{|s(t)|^2} = \lim_{P \rightarrow \infty} \frac{1}{2P} \int_{-P}^P |s(t)|^2 dt$$

die mittlere Leistung der Signalfunktion angibt. Durch

$$(2.3-15) \quad s_{\text{eff}} = \left(\lim_{P \rightarrow \infty} \frac{1}{2P} \int_{-P}^P |s(t)|^2 dt \right)^{\frac{1}{2}}$$

erhält man den Effektivwert oder quadratischen Mittelwert. Der zeitliche lineare Mittelwert des Betrages

$$(2.3-16) \quad s_F = \overline{|s(t)|} = \lim_{P \rightarrow \infty} \frac{1}{2P} \int_{-P}^P |s(t)| dt$$

heißt Flächenmittelwert.

Beispielsweise verschwindet für die Signalfunktion

$$s(t) = \sin(t)$$

der Gleichanteil (2.3-10), d.h. es ist

$$\overline{s} = \int_{-\pi}^{\pi} \sin(t) dt = 0$$

und damit

$$\hat{s} = \sin(t) .$$

Die mittlere Leistung (2.3-14) hat dann den Wert

$$\begin{aligned} \overline{|\sin(t)|^2} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |\sin^2(t)| dt = \\ &= \frac{4}{2\pi} \int_0^{\frac{\pi}{2}} \sin^2(t) dt = \frac{1}{2} \end{aligned}$$

und der Effektivwert (2.3-15) ist

$$s_{\text{eff}} = \sqrt{\frac{1}{2}} = \frac{\sqrt{2}}{2} .$$

Diese Differenz zwischen der maximalen Leistung (Musikleistung) und sog. effektiven Leistung (Sinusleistung) dürfte speziell Käufern von Stereoboxen und -verstärkern bzw. Musikinteressierten nicht ganz unbekannt sein.

Der Flächenmittelwert (2.3-16) ist in diesem Fall

$$\begin{aligned} s_F = \overline{|\sin(t)|} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |\sin(t)| dt = \\ &= 4 \int_0^{\frac{\pi}{2}} \sin(t) dt = 4 |-\cos(t)|_0^{\frac{\pi}{2}} = \frac{2}{\pi} . \end{aligned}$$

Die in (2.3-2) angegebene Signalfunktion läßt sich auch als Repräsentant eines stochastischen Prozesses (2.2-5) verstehen ^(2/11).

Bezeichnet die Hochzahl q wiederum den jeweiligen Repräsentanten, so wird man anstelle von (2.3-2) die Funktion in der Form

$$(2.3-17) \quad q_s = q_a \sin(\omega t + q_\psi) \quad \text{auffassen.}$$

Durch Superposition mehrerer Funktionen

$$(2.3-18) \quad \sum_{i=1}^n q_{a_i} \sin(\omega t + q_{\psi_i})$$

wird eine Folge von Zufallsvariablen und damit ein stochastischer Prozeß definiert, der als Modell für zahlreiche empirische Phänomene entscheidende praktische Bedeutung erlangt hat.

Für jeden Wert $t \in T$ stellt $s(t)$ also eine Zufallsgröße dar. Entspricht T einem Intervall auf der (reellen) Zahlengeraden

$$(2.3-19) \quad T = \{t_\ell \mid t_a \leq t_\ell \leq t_b\} \quad ,$$

spricht man von einem kontinuierlichen stochastischen Prozeß. Entspricht T hingegen einer abzählbaren Menge, z.B.

$$(2.3-20) \quad T = \{t_\ell \mid 0 \leq \ell \leq m\} \quad \text{mit } t_\ell < t_{\ell+1} \quad ,$$

handelt es sich um einen stochastischen Prozeß mit diskreter Zeit oder um einen diskreten stochastischen Prozeß.

Betrachtet man einen stochastischen Prozeß

$$\gamma(t) = \{q_s(t) \mid q=1,2,\dots\}$$

in einem speziellen Punkt t_1 , entsprechen die in t_1 von den Repräsentanten angenommenen Werte den Werten einer Zufallsvariablen

$$(2.3-21) \quad \gamma_{t_1} = \gamma(t_1) = [{}^1s(t_1), {}^2s(t_1), \dots] \quad .$$

Aus dieser Überlegung läßt sich als wesentlicher Begriff die Verteilungsfunktion $F_{t_1}(s)$ ableiten. Sie wird als Wahrscheinlichkeit p

definiert, mit der die Repräsentanten des Prozesses im Punkt t_1 Werte annehmen, die kleiner als der Argumentwert s_1 sind:

$$(2.3-22) \quad F_{t_1}(s_1) := p(s(t_1) < s_1).$$

Hieraus ergibt sich mit

$$(2.3-23) \quad p_{t_1}(s_1) := \frac{d F_{t_1}(s_1)}{d s_1}$$

die zugehörige Wahrscheinlichkeitsdichtefunktion, für die insbesondere gilt

$$(2.3-24) \quad \int_{-\infty}^{\infty} p_{t_1}(s) ds = 1.$$

Im Unterschied zur bisher beschriebenen zeitlichen Mittelung, bei der eine spezielle Funktion zugrundelag, bezieht sich die statistische Mittelung auf den stochastischen Prozeß selbst.

Die stochastische Mittelung führt zu den Erwartungswerten

$$(2.3-25) \quad E_{t_1}(s) := \int_{-\infty}^{\infty} s p_{t_1}(s) ds = \mu_{t_1}$$

und

$$(2.3-26) \quad E_{t_1}(s^2) := \int_{-\infty}^{\infty} s^2 p_{t_1}(s) ds.$$

Der Begriff der Varianz

$$(2.3-27) \quad \sigma_{t_1}^2 := E_{t_1}[(s-\mu)^2] = E_{t_1}(s^2) - \mu^2$$

bildet den Ausgangspunkt für eine Reihe wichtiger klassischer statistischer Kenngrößen und wird an späterer Stelle noch aufgegriffen.

Speziell bei einem harmonischen Prozeß (2.3-18) kann jeder Repräsentant aus sich selbst entstanden gedacht werden. Diese Eigen-

schaft, nach der alle den Prozeß charakterisierenden Wahrscheinlichkeitsdichtefunktionen invariant gegenüber beliebiger Translation des Parameters t sind, wird 'Stationarität' genannt.

Für eine beliebige Zeitverschiebung t_0 gilt also

$$(2.3-28) \quad p_{(t_q+t_0)}(q_s) = p_{t_q}(q_s) \quad (q=1,2,\dots) .$$

Insbesondere gilt für stationäre Prozesse

$$(2.3-29) \quad E(q_s) = \mu_q = \text{const.} .$$

Während bei stationären stochastischen Prozessen die zeitliche Mittelung konvergiert, ist dies nach dem Ergodentheorem (2/12) für nicht-stationäre Vorgänge nicht gesichert. Nur unter der Voraussetzung der Stationarität führen zeitliche und statistische Mittelung zum gleichen Ergebnis.

Obwohl die Annahme der Ergodizität in der Praxis höchst selten überprüfbar ist, setzt man sie zunächst voraus, um aus der zugänglichen Musterfunktion erste Rückschlüsse auf das Verhalten des interessierenden stochastischen Prozesses zu ziehen. Anhand der gewonnenen Ergebnisse wird dann beurteilt, in welchem Ausmaß diese Hypothese zutrifft oder nicht.

2.4 AUTO- UND KREUZKORRELATIONSFUNKTION

Den Ausdruck

$$(2.4-1) \quad \Omega_R = \overline{s_1(t) s_2(t)} = \lim_{P \rightarrow \infty} \frac{1}{2P} \int_{-P}^{+P} s_1(t) s_2(t) dt$$

zweier Signalfunktionen $s_1(t)$ und $s_2(t)$ bezeichnet man als Korrelationsfaktor. Für das feste Intervall ergibt sich

$$(2.4-2) \quad \Omega_R = \frac{1}{P_1 - P_2} \int_{P_2}^{P_1} s_1(t) s_2(t) dt .$$

Vorausgesetzt ist hier, daß der Gleichanteil (2.3-10) von $s_1(t)$ und $s_2(t)$ verschwindet.

Der Korrelationsfaktor ist maximal für

$$\Omega_{\max} = s_{\text{eff}_1} + s_{\text{eff}_2} ,$$

wobei s_{eff} gemäß (2.3-15) den jeweiligen Effektivwert meint. Die Normierung

$$(2.4-3) \quad \Omega(s_1(t), s_2(t)) = \Omega_{s_1, s_2} = \frac{\Omega_R}{\Omega_{\max}}$$

wird Korrelationskoeffizient genannt. Aus (2.4-3) folgt sofort, daß $-1 \leq \Omega_{s_1, s_2} \leq +1$ ist.

Als Kreuzkorrelationsfunktion zweier Funktionen $s_1(t)$, $s_2(t)$ wird der Ausdruck

$$(2.4-4) \quad \Omega_{KR}(\tau) = s_1(t)s_2(t+\tau) = \lim_{P \rightarrow \infty} \frac{1}{2P} \int_{-P}^{+P} s_1(t)s_2(t+\tau)dt$$

bzw.

$$(2.4-5) \quad \Omega_{KR}(\tau) = \frac{1}{P_1 - P_2} \int_{P_1}^{P_2} s_1(t)s_2(t+\tau)dt$$

verstanden. Da sich Schwankungen der Amplitude in der Regel störend auswirken, wird meist die normierte Kreuzkorrelationsfunktion

$$(2.4-6) \quad \Omega_N(\tau) = \frac{\Omega_{KR}(\tau)}{s_{\text{eff}_1} s_{\text{eff}_2}}$$

benutzt.

Ist nun $s_1(t) = s_2(t) = s(t)$, d.h. wird die Signalfunktion mit sich selbst korreliert, ergibt sich aus (2.4-4) die Autokorrelationsfunktion

$$(2.4-7) \quad \Omega_A(\tau) = \overline{s(t)s(t+\tau)} = \lim_{P \rightarrow \infty} \frac{1}{2P} \int_{-P}^{+P} s(t)s(t+\tau)dt ,$$

die für ein festes Intervall gemäß

$$(2.4-8) \quad \alpha_A(\tau) = \frac{1}{P_1 - P_2} \int_{P_1}^{P_2} s(t) s(t+\tau) dt$$

definiert wird.

Das Argument τ gibt wie auch sonst die Zeitverschiebung ('time lag') an. (2.4-7) bzw. (2.4-8) kann als ein Maß für die Übereinstimmung zwischen frühen und späteren Teilen und damit für die Prognostizierbarkeit einer Signalfunktion gelten.

Für rein stochastische Signale fällt $\alpha_A(\tau)$ mit wachsendem τ gegen Null ab, während sie für periodische Signalfunktionen im Punkt $\tau=P$ ein Maximum annimmt bzw. bei beliebig großem τ noch endliche Werte besitzt. In der Praxis wird die Berechnung von (2.4-8) oft auch als Methode zur Ermittlung der Periodenlänge einer Signalfunktion genutzt (2/13).

Im Unterschied zur obigen Definition der Korrelationsfunktion als zeitlicher Mittelwert liegt es bei stochastischen Prozessen (2.2-5) nahe, die Korrelationsfunktion im Sinne der stochastischen Mittelung zu definieren.

Betrachtet man die Zeitpunkte t_1 und t_2 mit $t_2 = t_1 + \tau$ und setzt $s_1 = s(t_1)$ und $s_2 = s(t_2)$ (d.h. die Variablen s_1 und s_2 stellen alle möglichen Werte der Repräsentanten des stochastischen Prozesses in den Punkten t_1 und t_2 dar), wird die Autokorrelationsfunktion eines stochastischen Prozesses durch den Erwartungswert

$$(2.4-9) \quad E(s_1, s_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s_1 s_2 p(s_1, s_2) ds_1 ds_2$$

bestimmt.

Hierbei ist $p(s_1, s_2)$ die Verbundwahrscheinlichkeitsdichtefunktion für die Zufallsgrößen s_1 und s_2 , die in Erweiterung von (2.3-23) durch

$$(2.4-10) \quad p(s_1, s_2) = \frac{d^2}{ds_1 ds_2} F(s_1, s_2) = \frac{d^2}{ds_1 ds_2} p(s(t_1) < s_1; s(t_2) < s_2)$$

festgelegt wird.

Unter der Bedingung der stochastischen Unabhängigkeit ist

$$(2.4-11) \quad p(s_1, s_2) = p_1(s_1) \cdot p_2(s_2) \quad .$$

Hierbei sind $p_1(s_1)$ und $p_2(s_2)$ die entsprechend (2.3-23) definierten eindimensionalen Dichtefunktionen. Die Bedingung (2.4-11) wird insbesondere für wachsendes τ zutreffen.

Für stationäre Vorgänge (2.3-28) soll voraussetzungsgemäß $p_1 = p_2$ sein. Während die Autokorrelationsfunktion (2.4-9) im allgemeinen von der Wahl der Punkte t_1 und t_2 bestimmt wird, ist sie für stationäre Prozesse nur von dem Abstand $\tau = t_2 - t_1$ abhängig.

Durch stochastische Mittelung läßt sich analog die Kreuzkorrelationsfunktion für stationäre Prozesse bestimmen. Sei für die Signalfunktion $s_1(t)$ in den Zeitpunkten t_1, t_2 mit $t_2 = t_1 + \tau$, $s_{11} = s_1(t_1)$, $s_{12} = s_1(t_2)$ und werde entsprechend für $s_2(t)$ gesetzt $s_{21} = s_2(t_1)$ sowie $s_{22} = s_2(t_2)$, dann kann die Kreuzkorrelationsfunktion als Erwartungswert notiert werden in der Form

$$(2.4-12) \quad E(s_{11}, s_{22}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s_{11} s_{22} p(s_{11}, s_{22}) ds_{11} ds_{22}$$

bzw.

$$(2.4-13) \quad E(s_{21}, s_{12}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s_{21} s_{12} p(s_{12}, s_{21}) ds_{12} ds_{21} \quad .$$

Gerade bei vielen Daten ist die Berechnung der Auto- und Kreuzkorrelation in der obigen Form relativ aufwendig. Im Punkt 3.4 wird aufbauend auf der Beschreibung eines effektiven Algorithmus kurz erläutert, auf welche Weise sich die erforderlichen numerischen Schritte erheblich vereinfachen lassen.

2.5 ANSÄTZE DER PARAMETEREXTRAKTION

2.5.1 EINFÜHRUNG

In einem ersten Auswertungsschritt wird man häufig die zu untersuchende Signalfunktion mittels der in den letzten Punkten angegebenen deskriptiven und statistischen Größen zu charakterisieren

suchen. Oft ist die so gewonnene und in speziellen Histogrammen aufbereitete Information jedoch unübersichtlich und noch zu wenig aussagekräftig, so daß man Merkmale oder Parameter extrahieren wird, die eine angemessenere Datenanalyse und -interpretation ermöglichen.

In Situationen, in denen aufgrund von Modellvorstellungen über das Verhalten der Signalfunktion bereits Vorkenntnisse existieren oder bestimmte Wahrscheinlichkeitsaussagen gemacht werden können, wird man einen analytischen Ausdruck wählen, der mit noch unbekannten Parametern die relevanten (funktionalen) Zusammenhänge wiedergibt (3/14).

Meist sind allerdings keine ausreichend explizierten Konzepte oder Modelle über den interessierenden psychophysiologischen Prozeß verfügbar und man wird sich für einen anderen Zugang der Extraktion hinreichend repräsentativer Merkmale entscheiden. Man wird mit möglichst wenigen Parametern die empirischen Funktionen hinreichend genau zu beschreiben versuchen, wobei der numerische Aufwand zur Ermittlung dieser Parameter in einem günstigen Verhältnis zu dem Vorteil zu stehen hat, der durch die Verringerung der Datenzahl erreicht worden ist.

Aus diesen Überlegungen heraus ist es naheliegend zu versuchen, die gegebene Signalfunktion $s(t)$ durch numerisch möglichst einfach auswertbare Funktionen wie etwa Polynome, rationale Funktionen oder speziell trigonometrische Funktionen anzunähern.

Oft wird für die approximierende Funktion $g(t)$ ein Polynom verwandt. Zum einen kann durch Polynome beliebig hohen Grades jede stetige Funktion beliebig gut angenähert werden (2/15), zum anderen lassen sich Polynome numerisch leicht auswerten.

Um nun die Güte der Näherung von $g(t)$ an $s(t)$ ermitteln zu können, wird man eine geeignete Gütefunktion oder "Norm" definieren. Auf diesen Aspekt wird unter Punkt (2.5.2.1) Bezug genommen.

In (2.5.2.2) wird aufgezeigt, daß eine numerisch besonders einfache Lösung des Approximationsproblems dann zu erzielen ist, wenn das Funktionensystem orthogonal und normiert ist.

Wird zur Approximation ein derartiges orthonormales Funktionensystem verwandt, ergeben sich zwei unterschiedliche Fälle. Einmal kann man das Orthonormalsystem fest vorgeben und $s(t)$ hinsichtlich der benutzten Norm möglichst gut anzupassen versuchen. Andererseits kann man aus den zahlreichen orthonormalen Funktionensystemen jenes auswählen, für das der Fehler der Näherung minimal wird. Man paßt hier also das Funktionensystem den vorgegebenen Daten möglichst gut an und spricht daher von problemabhängiger Approximation oder - da es sich bei diesen Approximationen um endliche oder unendliche Linearkombinationen orthonormaler Funktionen und somit um Summen handelt - auch von problemabhängiger Reihenentwicklung. Demzufolge handelt es sich im ersten Fall um eine sog. problemunabhängige Approximation.

Die praktisch relevanten Funktionen, die zur problemunabhängigen Approximation benutzt werden, finden sich in Punkt (2.5.3). In (2.5.4) wird das Prinzip der problemabhängigen Wahl eines angemessenen Funktionensystems kurz erläutert.

Aus den bisherigen Ausführungen ist deutlich geworden, daß die Approximation von $g(t)$ an $s(t)$ eine Methode zur Reduktion der Daten und damit der Merkmalsextraktion ist, die daher auch in der Praxis relativ verbreitet ist. Demgegenüber ist ein weiterer Ansatz speziell in der Psychologie vergleichsweise unbekannt, der insbesondere dann angezeigt ist, wenn die zu approximierende Funktion in bestimmten Teilintervallen oder Segmenten eine stark verändernde Gestalt besitzt und numerisch durch bekannte Basisfunktionen nur schwer zu handhaben ist. Dieser Ansatz, bei dem eine möglichst glatte Approximation angestrebt wird und der als einen Kompromiß zwischen Polygonzug und Interpolationspolynom verstanden werden kann, wird einführend in Punkt (2.5.5) aufgegriffen.

2.5.2 ALLGEMEINE VORAUSSETZUNGEN

2.5.2.1 NORM

Eine Funktion $\|\cdot\|$ von einer Menge (z.B. von Funktionen) G in die nicht-negativen reellen Zahlen $\|\cdot\| : G \rightarrow \mathbb{R}_+$ wird als Norm bezeichnet, wenn für zwei beliebige Elemente $s \in G$ und $g \in G$ und $\alpha \in \mathbb{R}$ die Bedingungen

$$(2.5.2.1-1) \quad \|s\| \geq 0$$

$$(2.5.2.1-2) \quad \|s\| = 0 \Leftrightarrow s = 0$$

$$(2.5.2.1-3) \quad \|\alpha s\| = |\alpha| \|s\|$$

$$(2.5.2.1-4) \quad \|s+g\| \leq \|s\| + \|g\|$$

erfüllt sind ^(2/16). Man bezeichnet eine Menge (z.B. von Funktionen), auf denen eine Normfunktion $\|\cdot\|$ mit den Eigenschaften (2.4.2.1.1-4) definiert ist, auch als normierten Raum. Ein linearer Raum mit Norm heißt auch linearer normierter Raum. Im folgenden wollen wir diese Menge G stets als Funktionenraum von Funktionen $s(t)$, $g(t)$, $h(t)$, ... auffassen.

Eine ganze Klasse von Normen wird durch die L_r -Norm festgelegt ($r > 0$)

$$(2.5.2.1-5) \quad L_r(h) := \|h\|_r := \left[\int_{-\infty}^{\infty} |h(t)|^r dt \right]^{\frac{1}{r}}.$$

Von praktischer Relevanz sind vor allem die Fälle $r=1$, $r=2$ und $r=\infty$.

Die L_∞ -Norm, TSCHEBYSCHEV- oder auch Minimax-Norm

$$(2.5.2.1-6) \quad L_\infty(h) := \|h\|_\infty = \max_{0 \leq t < \infty} |h(t)|$$

zielt auf die Minimierung des maximalen Fehlerbetrages.

Legt man die L_2 -Norm bzw. ihr Quadrat die Fehlerfunktion

$$(2.5.2.1-7) \quad [L_2(h)]^2 = \int_{-\infty}^{\infty} h(t)^2 dt$$

zugrunde, führt dies zur bekannten Ausgleichsrechnung nach der Kleinst-Quadrat Methode, welche wegen ihrer für empirische Daten vorteilhaften Eigenschaften in nahezu allen wissenschaftlichen Bereichen Anwendung findet.

Der Effekt des Parameters r in (2.5.2.1-5) ist darin zu sehen, daß mit wachsendem r Bereiche, in denen Fehler auftauchen, stärker bewertet werden. Man wird daher je nach den Erfordernissen der Untersuchungsfragestellung unterschiedliche Werte für r einsetzen.

Erscheint desweiteren eine Gewichtung der durch die Norm ermittelten Abweichungen in verschiedenen Teilintervallen angemessen, wird man (2.5.2.1-5) durch eine positive Funktion v gewichten. Mit $v(t) > 0$ und $(t \in \mathbb{R})$ ergibt sich dann

$$(2.5.2.1-8) \quad L_r(h, v) = \left[\int_{-\infty}^{\infty} v(t) |h(t)|^r dt \right]^{1/r}.$$

Häufig wird im Intervall $[0, 1]$

$$v(t) = t^\alpha (1-t)^\beta \quad \text{mit } \alpha, \beta \geq -1/2 \text{ gesetzt.}$$

So wird etwa $v(t) = [t(1-t)]^{-1/2}$ benutzt, wenn den Intervallenden eine größere Bedeutung zukommen soll, wenn das spezielle Approximationsproblem beispielsweise darauf zielt, daß an den Intervallenden besonders gut approximiert werden soll.

Eine andere Generalisation der L_r -Norm stellt dieser als SOBOLEW-Norm bezeichnete Ausdruck dar.

$$(2.5.2.1-9) \quad \|h\|_{H_m^r} := \left[\sum_{i=0}^m \int_{-\infty}^{\infty} \left| \frac{d^i h}{dt^i} \right|^r dt \right]^{1/r}$$

SOBOLEW-Normen werden zur Beurteilung der Glätte einer Funktion benutzt, in solchen Fällen also, in denen nicht nur die Funktion selbst sondern auch die ersten m Ableitungen simultan zu approximieren sind.

2.5.2.2 ORTHONORMALSYSTEM

Sei H ein linearer Raum, in dem ein Skalarprodukt gegeben ist, d.h. für je zwei Elemente $s_i, s_j \in H$ gibt es eine reellwertige Funktion (s_i, s_j) mit den Eigenschaften

$$(2.5.2.2-1) \quad (s_i, s_j) = (s_j, s_i)$$

$$(2.5.2.2-2) \quad (\alpha s_i, s_j) = \alpha (s_i, s_j) \quad (\alpha \in \mathbb{R})$$

$$(2.5.2.2-3) \quad (s_k + s_i, s_j) = (s_k, s_j) + (s_i, s_j)$$

$$(2.5.2.2-4) \quad (s_i, s_i) \geq 0, \quad (s_i, s_i) = 0 \Leftrightarrow s_i = 0.$$

Durch dieses Skalarprodukt oder auch "inneres Produkt" ist eine Norm $\| \cdot \|$ auf H gemäß

$$(2.5.2.2-5) \quad \|s\| = [(s,s)]^{1/2}$$

gegeben. Einen solchen Raum nennt man Skalarprodukt-Raum. Wegen (2.5.2.2-5) ist er immer ein normierter Raum. Ist dieser Raum bezüglich der unter (2.5.2.2-5) gegebenen Norm vollständig, d.h. jede Cauchyfolge konvergiert gegen ein Element aus H ^(2/17), so nennt man H auch Hilbert-Raum.

Beispielsweise wird der m -dimensionale euklidische Raum $\mathbb{R}^m$ zum Hilbert-Raum, wenn man das Skalarprodukt zwischen

$$s_i = \{\sigma_1^i, \dots, \sigma_m^i\} \quad \text{und}$$

$$s_j = \{\sigma_1^j, \dots, \sigma_m^j\}$$

durch

$$(s_i, s_j) = \sum_{k=1}^m \sigma_k^i \sigma_k^j$$

definiert. Für dieses Skalarprodukt ergibt die Norm (2.5.2.2-5) die bekannte euklidische Norm.

Für ein Skalarprodukt ergibt sich aus (2.5.2.2-1) - (2.5.2.2-4) die SCHWARZ'sche Ungleichung

$$(2.5.2.2-6) \quad |(s_i, s_j)| \leq [(s_i, s_i)(s_j, s_j)]^{1/2}.$$

Aufbauend auf diesen Festsetzungen kann das bisher allgemein formulierte Approximationsproblem aus (2.5.1) wie folgt für den HILBERT-Raum angegeben werden.

Sei H ein HILBERT-Raum und $H_u \subset H$ ein m -dimensionaler Unterraum. Dann existieren in H_u m linear unabhängige normierte Elemente $h_1, h_2, \dots, h_m$ derart, daß sich jedes Element $h \in H_u$ eindeutig als Linearkombination

$$(2.5.2.2-7) \quad h = \sum_{k=1}^m \alpha_k h_k \quad \text{mit} \quad \alpha_k \in \mathbb{R}$$

darstellen läßt.

Da in linear normierten Räumen für ein beliebiges $s \in H$, $s \notin H_U$ genau eine beste Approximation $h \in H_U$ zu s existiert (2/18), besteht das eigentliche Problem darin, die Koeffizienten α_k für dieses h zu ermitteln. Gesucht ist also ein $h \in H_U$ mit $\|s-h\| \leq \|s-y\|$ für alle $y \in H_U$.

Anstelle der Norm

$$(2.5.2.2-8) \quad \|s-h\| = \left\| s - \sum_{k=1}^m \alpha_k h_k \right\| = \min_{y \in H} \|s-y\|$$

werde die quadratische Form der Norm

$$(2.5.2.2-9) \quad \|s-h\|^2 = \|s\|^2 - 2 \sum_{k=1}^m \alpha_k (s, h_k) + \sum_{\ell, k=1}^m \alpha_\ell \alpha_k (h_\ell, h_k)$$

zugrundegelegt (2/19), welche also eine quadratische Funktion der m reellen Koeffizienten ist. (2.5.2.2-9) nimmt ein Minimum an, wenn die ersten partiellen Ableitungen nach den α_k verschwinden, d.h.

$$(2.5.2.2-10) \quad (s, h_\ell) - \sum_{k=1}^m \alpha_k (h_k, h_\ell) = 0 \quad (1 \leq \ell \leq m)$$

bzw. in abgekürzter Schreibweise

$$(2.5.2.2-11) \quad (s-h, h_\ell) = 0 \quad (1 \leq \ell \leq m).$$

Damit erhält man die gesuchten Koeffizienten durch Lösung des linearen Gleichungssystems

$$(2.5.2.2-12) \quad \sum_{k=1}^m \alpha_k (h_k, h_\ell) = (s, h_\ell) \quad (1 \leq \ell \leq m).$$

Wie man der Gleichung (2.5.2.2-12) ansieht, ist die Lösung dann besonders einfach zu berechnen und daher praktisch wichtig, wenn die Elemente $h_1, \dots, h_m$ ein Orthonormalsystem bilden, wenn also gilt

$$(2.5.2.2-13) \quad (h_k, h_\ell) = \delta_{k\ell} = \begin{cases} 1 & \ell=k \\ 0 & \text{sonst} \end{cases} .$$

In diesem Fall kann man die Lösung unmittelbar ablesen

$$(2.5.2.2-14) \quad \alpha_k = (s, h_k) .$$

Ein weiterer in der Praxis wichtiger Vorteil der Benutzung von orthogonalen Systemen besteht darin, daß man bei Hinzunahme weiterer Basiselemente (Vergrößerung des Raumes H_u) nur die weiteren Koeffizienten neu berechnen muß. Eine einmal ermittelte Lösung mit $m' < m$ Basiselementen stellt m.a.W. bereits die beste m' -dimensionale Approximation auch im m -dimensionalen Raum dar.

Ein vollständiges ONS (VONS) ist ein ONS von Elementen $y_1, y_2, \dots, y_H$, so daß jedes $y \in H$ darstellbar ist

$$(2.5.2.2-15) \quad y = \sum_{i=1}^m \alpha_i y_i \quad \text{mit}$$

$$(2.5.2.2-16) \quad \alpha_i = (y_i, y) .$$

Die α_i in (2.5.2.2-16) werden auch als Fourierkoeffizienten bezeichnet. (2.5.2.2-15) gilt im übrigen auch für $m \rightarrow \infty$.

Eine hinreichende und notwendige Bedingung für ein ONS, vollständig zu sein, ist das Erfülltsein der Vollständigkeitsrelation

$$(2.5.2.2-17) \quad \text{Aus } s \in H \text{ mit } (s, y_i) = 0 \text{ für alle } i \Rightarrow s = 0 \quad (3/20)$$

2.5.3 PROBLEMUNABHÄNGIGE PARAMETEREXTRAKTION

Auf der Basis der bisherigen Ausführungen können einige der in der Praxis wichtigen orthogonalen Funktionensysteme wie folgt kurz angegeben werden. Ihre Eigenschaften lassen sich aus den obigen Darlegungen ebenso einfach ersehen.

Von praktischer Relevanz dürfte insbesondere sein, daß man das Funktionensystem nach speziellen Kriterien auswählen wird, die bezüglich der Struktur bzw. des Verhaltens der zu untersuchenden Daten als angemessen ausgewiesen werden können. Sehr häufig werden daher die trigonometrischen Funktionen (2.5.3.1) herangezogen, deren Koeffizienten zudem anschaulich interpretierbar sind ^(2/21). Da man jedoch nicht selten die berechneten Koeffizienten als neue Merkmale für eine weitere statistische Analyse verwendet, eignen sich ebenso andere, unter Umständen der Datenqualität eher entsprechende und numerisch leichter auszuwertende Funktionensysteme zur Extraktion der gesuchten Parameter. Aus dieser Sicht dürften das in Punkt 2.5.3.2 beschriebene System der Walsh-Funktionen oder auch das in 2.5.3.4 skizzierte Orthogonalsystem der Tschebyscheffschen Polynome in der Praxis zu interessanten Anwendungen führen.

2.5.3.1 TRIGONOMETRISCHE FUNKTIONEN

Das bekannteste System orthogonaler und normierter Funktionen dürften die trigonometrischen Funktionen bilden, was u.a. darauf zurückzuführen sein dürfte, daß die sin-/cos-Funktion als Idealtyp zahlreicher empirischer Prozesse angesehen werden können und sich die berechneten Koeffizienten im Sinne neu gewonnener Parameter (physikalisch) anschaulich interpretieren lassen.

Sei $H := L_2[0, 2\pi]$ und das Skalarprodukt durch

$$(s_i, s_j) := \int_0^{2\pi} s_i(t) s_j(t) dt \quad \text{gegeben.}$$

Die Funktionen

$$\frac{1}{\sqrt{2\pi}}, \quad \frac{\cos t}{\sqrt{\pi}}, \quad \frac{\cos 2t}{\sqrt{\pi}}, \quad \dots; \quad \frac{\sin t}{\sqrt{\pi}}, \quad \frac{\sin 2t}{\sqrt{\pi}}, \quad \dots$$

bilden ein VONS. Es gilt für die Signalfunktion $s(t) \in L_2[0, 2\pi]$

$$(2.5.3.1-1) \quad s(t) = \frac{a_0}{2} + \sum_{i=1}^{\infty} (a_i \cos it + b_i \sin it),$$

wobei sich die Koeffizienten wegen (2.5.2.2-14) gemäß

$$(2.5.3.1-2) \quad a_0 = \frac{1}{\pi} \int_0^{2\pi} s(t) dt$$

$$(2.5.3.1-3) \quad a_i = \frac{1}{\pi} \int_0^{2\pi} s(t) \cos it \, dt \quad \text{und}$$

$$(2.5.3.1-4) \quad b_i = \frac{1}{\pi} \int_0^{2\pi} s(t) \sin it \, dt \quad (1 \leq i < \infty)$$

berechnen. Für (2.5.3.1-1) ist mit

$$e^{j it} = \cos it + j \sin it \quad (j := \sqrt{-1})$$

(Eulersche Gleichung) auch die numerisch vorteilhaftere komplexe Schreibweise üblich

$$(2.5.3.1-5) \quad s(t) = \sum_{i=-\infty}^{\infty} \alpha_i e^{j it} \quad (i = 0, \pm 1, \pm 2, \dots)$$

mit

$$\alpha_i = \frac{1}{2\pi} \int_{-\pi}^{\pi} s(t) e^{-j it} \, dt \quad .$$

2.5.3.2 WALSH-FUNKTIONEN

Zunehmende Bedeutung erlangen in neuerer Zeit die Mäander- oder auch Walsh-Funktionen. Sie bilden ein orthogonales und vollständiges Funktionensystem, sind jedoch im Unterschied zu den trigonometrischen Funktionen nicht stetig.

Die Walsh-Funktionen können rekursiv in $[-1, 1]$ durch

$$(2.5.3.2-1) \quad \text{wal}(0, t) = \begin{cases} 1 & \text{für } -1/2 \leq t < 1/2 \\ 0 & \text{für } |t| > 1/2, \quad t = 1/2 \end{cases} \quad ,$$

$$\begin{aligned} \text{wal}(2v+\alpha, t) &= (-1)^{[v/2]+\alpha} [\text{wal}(v, 2t+1/2) + \\ &\quad (-1)^{v+\alpha} \text{wal}(v, 2t-1/2)] \end{aligned}$$

mit

$$[v/2] = \text{Entier } (v/2),$$

$$v = 0, 1, \dots, n-1 \quad \text{und}$$

$$\alpha = 0, 1$$

definiert werden.

In Analogie zum Begriff der Frequenz wird unter der Größe v die Sequenz verstanden und in Zeichenwechsel pro vorgegebenen Zeitintervall geteilt angegeben $(2/22)$.

Gerade Walsh-Funktionen werden von ungeraden gemäß

$$(2.5.3.2-2) \quad \text{wal}(2i, t) = \text{cal}(i, t) \quad (1 \leq i < \infty)$$

$$(2.5.3.2-3) \quad \text{wal}(2i-1, t) = \text{sal}(i, t)$$

unterschieden. Das Skalarprodukt in $L_2[0, T]$ mit $T = 2\pi$

$$(s_i, s_j) = \int_0^T s_i s_j \, dt$$

reduziert sich zu Summen, da die Walsh-Funktionen konstant sind. Die Eigenschaft der Orthogonalität ergibt sich aus

$$\int_t \text{wal}(i, t) \text{wal}(j, t) = n \delta_{ij} \quad (n \in \mathbb{N})$$

und die der Vollständigkeit aus

$$\sum_i \text{wal}(i, t) \text{wal}(i, t') = n \delta_{tt'}$$

In Entsprechung zu (2.5.3.1-1) kann die Walsh-Reihe durch

$$s(t) = a_0 \text{wal}(0, t) + \sum_{i=1}^{\infty} a_i \text{wal}(i, t)$$

bzw. unter Bezug auf die Unterscheidung (2.5.3.2-2/3) durch

$$(2.5.3.2-4) \quad s(t) = a_0 \text{ wal}(0,t) + \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_i \text{ sal}(i,t) + b_j \text{ cal}(j,t)$$

mit

$$(2.5.3.2-5) \quad a_0 = \frac{2}{T} \int_0^T s(t) \text{ wal}(0,t) dt$$

$$(2.5.3.2-6) \quad a_i = \frac{1}{T} \int_0^T s(t) \text{ sal}(i,t) dt$$

$$(2.5.3.2-7) \quad b_j = \frac{1}{T} \int_0^T s(t) \text{ cal}(j,t) dt$$

angegeben werden.

2.5.3.3 BEZIEHUNG ZWISCHEN DEN TRIGONOMETRISCHEN UND WALSH-FUNKTIONEN

Der Zusammenhang zwischen der Fourier- und Walsh-Reihe ergibt sich bereits durch elementare Betrachtungen ^(2/23). Da die Funktion $\text{wal}(0,t)$ definitionsgemäß den Wert 1 besitzt, erhält man für a_0 in (2.5.3.2-5) den gleichen Wert wie in (2.5.3.1-2). Bildet man weiterhin den Grenzwert für (2.5.3.2-4), ergibt sich sofort, daß beide Reihen zu dem gleichen Ergebnis führen.

Setzt man also Gleichung (2.5.3.2-4) für $i, j \rightarrow \infty$ in die Bestimmungsgleichung der cos-Koeffizienten (2.5.3.1-3) ein, ergibt sich

$$(2.5.3.3-1) \quad a_k = \frac{2}{T} \int_0^T \left[\frac{a_0}{T} + \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_i \text{ sal}(i,t) + b_j \text{ cal}(j,t) \right] \cos k\omega t dt$$

und nach einfacher Umformung

$$(2.5.3.3-2) \quad a_k = a_0 + \frac{2}{T} \sum_{i=1}^{\infty} a_i \left[\int_0^T \text{sal}(i,t) \cos k\omega t dt \right] +$$

$$\frac{2}{T} \sum_{j=1}^{\infty} b_j \left[\int_0^T \text{cal}(j,t) \cos k\omega t dt \right] .$$

Hierbei entsprechen die Ausdrücke in den eckigen Klammern den Fourier-Koeffizienten für $\text{sal}(i,t)$ und $\text{cal}(j,t)$. Ersetzt man beide Terme durch $a_i(\text{sal})$ und $b_j(\text{cal})$, kann man einfacher schreiben

$$(2.5.3.3-3) \quad a_i = a_0 + \frac{2}{T} \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} a_k a_i(\text{sal}) + b_j a_i(\text{cal}) .$$

Analog lassen sich die Koeffizienten b_i bestimmen

$$(2.5.3.3-4) \quad b_i = a_0 + \frac{2}{T} \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} a_k b_i(\text{sal}) + b_j b_i(\text{cal}) .$$

Will man andererseits die Fourier-Reihe durch Terme der Walsh-Reihe ausdrücken, wird man lediglich die in (2.5.3.3-3) und (2.5.3.3-4) definierten Koeffizienten in die Grundgleichung (2.5.3.1-1) einsetzen.

$$(2.5.3.3-5) \quad s(t) = \frac{a_0}{2} + \sum_{i=1}^{\infty} \left\{ \left[a_0 + \frac{2}{T} \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} (a_k a_i(\text{sal}) + b_j a_i(\text{cal})) \right] \cos i\omega t \right. \\ \left. + \left[a_0 + \frac{2}{T} \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} (a_k b_i(\text{sal}) + b_j b_i(\text{cal})) \right] \sin i\omega t \right\}$$

Wegen der Orthogonalität der Walsh- und Fourier-Funktionen vereinfacht sich im übrigen (2.5.3.3-5) zu

$$s(t) = \frac{a_0}{2} + \frac{2}{T} \sum_{i=1}^{\infty} \sum_{\ell=1}^{\infty} \sum_{j=1}^{\infty} (b_j a_i(\text{cal}) \cos i\omega t + a_{\ell} b_i(\text{sal}) \sin i\omega t) \quad (2/24).$$

Man wird m.a.W. in der Regel beliebige stetige Signalfunktionen angemessener durch eine Fourierreihe annähern können, während bei diskreten Funktionen mit rechteck- oder pulsartigem Charakter eine Entwicklung nach den Walsh-Funktionen vorteilhafter ist.

2.5.3.4 ANDERE ORTHOGONALE FUNKTIONEN

Während insbesondere die trigonometrischen Funktionen zur Parameterextraktion relativ häufig in der Praxis Verwendung finden, werden die nachfolgend beispielhaft angegebenen anderen Systeme orthogonaler Funktionen vergleichsweise selten benutzt. Angesichts der fortschreitenden Formalisierung der Wissenschaften dürften sie jedoch steigende Bedeutung besitzen, zumal sie bei spezifischen Anwendungsproblemen eine zweckmäßige und angemessene Lösung ermöglichen.

Sei $H := C[a, b]$ und das Skalarprodukt

$$(s_i, s_j) = \int_a^b s_i(t) s_j(t) p(t) dt$$

mit $p(t) > 0$ als "Belegungsfunktion" (2/25).

Durch Orthogonalisierung (z.B. mit Hilfe des Schmidtschen Verfahrens) der Funktionen

$$\sqrt{p(t)}, t\sqrt{p(t)}, t^2\sqrt{p(t)}, \dots$$

im Intervall $[a, b]$ lassen sich Orthogonalsysteme von Funktionen erzeugen. Die zugehörigen Faktoren von $p(t)$ sind dann Polynome $Q_0(t), Q_1(t), Q_2(t), \dots$ vom Grad $0, 1, 2, \dots$ und können als "die zur Belegung $p(t)$ gehörigen Polynome" (2/26) aufgefaßt werden. Die Vorgehensweise entspricht m.a.W. den in den vorangegangenen Punkten aufgezeigten Schritten mit dem Unterschied, daß in diesem Fall eine spezielle Funktionenklasse (Polynome) und spezielle Skalarprodukte betrachtet werden. Nach Orthogonalisierung der Basis lassen sich die betreffenden Koeffizienten dann gemäß (2.5.2.2-14) berechnen.

So ergeben sich etwa für

$$(2.5.3.4-1) \quad p(t) = 1, \quad a = -1 \quad \text{und} \quad b = 1$$

die Legendreschen Polynome $P_n(t)$.

Auch in der Psychophysiologie nicht unbekannt sind die Tschebyscheffschen Polynome

$$(2.5.3.4-2) \quad T_n(t) = \frac{1}{2^{n-1}} \cos(n \arccos t)$$

mit

$$p(t) = \frac{1}{\sqrt{1-t^2}}, \quad a = -1 \text{ und } b = 1.$$

Man erhält durch

$$(2.5.3.4-3) \quad p(t) = e^{-t^2}, \quad a = -\infty \text{ und } b = \infty$$

die Hermitschen Polynome, sowie durch

$$(2.5.3.4-4) \quad p(t) = e^{-t}, \quad a = 0 \text{ und } b = \infty$$

die Laguerreschen Polynome.

Schließlich führt

$$(2.5.3.4-5) \quad p(t) = (1-t)^p (1+t)^q \quad (p, q > -1) \quad \text{mit } a = 0, b = 1$$

auf die Jacobschen oder hypergeometrischen Polynome.

Da man jede Basis orthonormalisieren kann und es unendlich viele positive Gewichtsfunktionen gibt, ließe sich die Reihe solcher Beispiele noch beliebig erweitern.

2.5.4 PROBLEMAPABHÄNGIGE PARAMETEREXTRAKTION

Im Unterschied zur bisher betrachteten problemunabhängigen Parameterextraktion, bei der das Orthonormalsystem fest vorgegeben worden ist und die empirische Funktion hinsichtlich der benutzten Norm möglichst gut anzupassen ist, wird bei der problemabhängigen Parameterextraktion aus den zahlreichen orthonormalen Funktionensystemen jenes ausgewählt, für das der Fehler gemäß der geeignet bestimmten

Norm minimal wird. Das Prinzip dieser Vorgehensweise soll hier nur kurz umrissen werden, wobei wir uns auf die Approximation durch n -dimensionale Vektoren, d.h. den $\mathbb{R}^n$ beschränken.

Es werde vereinfachend angenommen, daß die erhobenen Meßwerte zuvor, etwa über

$$z_{ij} = (s_{ij} - \bar{s}_{.j}) / \sigma_j$$

mit σ_j als Standardabweichung und $\bar{s}_{.j}$ als Mittelwert des j -ten Merkmals, standardisiert und in einer Datenmatrix Z zusammengefaßt worden sind (2/27).

Bezeichne Z die standardisierte (n,m) -Datenmatrix, F das gesuchte (n,r) mit $k \leq m$ Orthonormalsystem und A die (m,r) -Matrix der Gewichtungsfaktoren, dann soll gelten

$$(2.5.4-1) \quad ||| Z - F A' |||_E^2 \rightarrow \text{Min}.$$

Statt der meist benutzten Euklidischen Norm kann ebenso eine etwa der in (2.5.2.1) angegebenen Normen als Fehlermaß dienen. Für (2.5.4-1) erhält man in Komponentenschreibweise

$$(2.5.4-2) \quad Q = \sum_{i=1}^n \sum_{\ell=1}^m (z_{i\ell} - \sum_{j=1}^r \alpha_{j\ell} f_{ij})^2 \rightarrow \text{Min}.$$

Der Index m gibt die Zahl der Elemente (Signalfunktionen) und n die Zahl der gemessenen Datenpunkte oder Merkmale an, während r die Zahl der Parameter (Orthonormalvektoren) meint ($1 \leq j \leq r$; $1 \leq \ell \leq m$; $1 \leq i \leq n$ und $r \leq m$).

Eine Lösung erhält man nach Sätzen aus der Linearen Algebra (Courantsches Minimax-Theorem) ^(2/28) durch Lösung des Eigenwertproblems

$$(2.5.4-3) \quad R U = U \Lambda,$$

wobei

$$R = \frac{1}{n} Z' Z$$

eine symmetrische positiv definite Matrix ist und die Spalten von U aus den zu den r größten Eigenwerten gehörenden Eigenvektoren gebildet werden. Die Diagonalmatrix $\Lambda = \text{diag}(\lambda_i)$ enthält die (positiven) Eigenwerte von R. Es ergibt sich als Lösung:

$$(2.5.4-4) \quad A = U \text{diag}(\lambda_i^{1/2}) \quad \text{und}$$

$$F = Z U \text{diag}(\lambda_i^{-1/2}) .$$

Die Spalten von F bilden also das System orthonormaler Vektoren. Die zugehörigen Linearkombinationen werden in der Literatur häufig unter der Bezeichnung "Karhunen-Loeve Entwicklung" angegeben (2/29). Eine Verallgemeinerung des obigen Prinzips auf unendliche Dimensionen ist naheliegend.

Der Zusammenhang zur Faktoren- bzw. Hauptkomponentenanalyse ist im übrigen unschwer zu erkennen, als A der Matrix der (unrotierten) Faktorenladungen und F der Matrix der zugehörigen Faktorenwerte oder Hauptkomponenten entspricht.

2.5.5 SEGMENTWEISE PARAMETEREXTRAKTION

2.5.5.1 EINFÜHRUNG

In den bisherigen Ausführungen wurde unterstellt, daß die zu approximierende Funktion im gesamten vorgegebenen Intervall einen annähernd gleichen oder zumindest vergleichbaren Verlauf aufweist und etwa durch ein orthonormales Funktionensystem oder eine andere numerisch handhabbare Funktion gut angenähert werden kann.

Besitzt die zu approximierende Funktion in bestimmten Teilintervallen eine sich sehr verändernde Gestalt oder oszilliert sie stark, was durch Meßungenauigkeiten, andere empirische Fehlerquellen oder durch den untersuchten psychophysiologischen Prozeß selbst bedingt sein kann, wird man geeignete Teilbereiche oder Segmente zu konstruieren versuchen und in jedem Segment die Funktion $s(t)$ gesondert betrachten.

Die hierzu erforderliche Zerlegung des Gesamtintervalls $[a, b]$, die durch ihre Übergangsstellen oder Knoten u_i charakterisiert ist,

$$a = u_0 < u_1 < \dots < u_j < \dots < u_{r+1} = b$$

wird man je nach praktischen Aspekten unterschiedlich vornehmen. Man wird jedoch allgemein fordern, daß an den Stützstellen die approximierenden Funktionen selbst und gewisse Ableitungen möglichst ohne Sprungstellen miteinander verbunden sind.

Segmentweise zusammengesetzte Funktionen, die an den Übergangsstellen glatt verheftet sind, werden in der Mathematik unter die Bezeichnung "Spline-Funktionen" gefaßt und bieten sich aufgrund ihrer Eigenschaften zur möglichst glatten Approximation empirisch gewonnener und numerisch ansonsten schwer auswertbarer Kurven besonders an.

So entstammt die Bezeichnung "Spline" auch einem praktischen Kontext. Im Schiffsbau wurde früher eine Straklatte, die im Englischen "Spline" heißt, zur Befestigung der in Längsrichtung des Schiffes verlaufenden Planken verwandt. In Anlehnung daran benutzten technische Zeichner einen dünnen metallischen Stab (Spline), um eine möglichst glatte Kurve zu zeichnen. Hierzu wurde der Spline in jenen Punkten fixiert, durch welche die Kurve verlaufen sollte.

Die Forderung einer möglichst glatten Approximation führt mathematisch gesehen zu Funktionen, die (zweimal) stetig differenzierbar sind und in Teilintervallen mit Polynomen (z.B. dritten Grades) übereinstimmen. Diese praktisch wohl wichtigsten Spline-Funktionen vom Grad $n=3$ werden Kubische Spline-Funktionen genannt und nachfolgend kurz beschrieben.

Verständlicherweise führt das Problem der segmentweisen Approximation auf das Problem der Interpolation mit Spline-Funktion, das entsprechend formuliert werden kann.

2.5.2.2 KUBISCHE SPLINES

Unter Spline-Funktionen können solche Funktionen verstanden werden, die aus gewissen Basisfunktionen, in der Regel Polynome n -ten

Grades stückweise zusammengesetzt sind in der Form, daß sie schließlich $(n-1)$ mal stetig differenzierbar sind ^(2/30). Die Forderung der $(n-1)$ mal stetigen Differenzierbarkeit von Spline-Funktionen zielt darauf, daß sie an den Eckpunkten der Teilintervalle genügend glatt von einem Polynom zum anderen übergehen.

Für die Spline-Funktion $c(t)$ n -ter Ordnung gilt also

$$(2.5.5.2-1) \quad c(t) \in C^{n-1} [a, b]$$

mit C^{n-1} als Menge der in $[a, b]$ $(n-1)$ mal stetig differenzierbaren Funktionen.

Von praktischer Relevanz sind insbesondere Spline-Funktionen dritten Grades oder kurz kubische Splines, die in der Form

$$(2.5.5.2-2) \quad c(t) = \sum_{j=1}^r p_j(t) = a_j + b_j t + c_j t^2 + d_j t^3$$

$$\text{für } u_j \leq t \leq u_{j+1} \quad (0 \leq j \leq r)$$

$$\text{mit } u_{j+1} = u_j + \delta(u_j)$$

notiert werden können. $\delta(u_j)$ bezeichnet den Abstand zwischen den Stützstellen oder Knoten. Oft ist vereinfachend

$$(2.5.5.2-3) \quad \delta(u_j) = \delta(u_{j+1}) = \text{konst.} \quad .$$

Gefordert wird Stetigkeit bis zur zweiten Ableitung

$$(2.5.5.2-4) \quad p_j^{(k)}(u_j) = p_{j+1}^{(k)}(u_j) \quad \begin{matrix} (0 \leq k \leq 2) \\ (1 \leq j \leq r) \end{matrix} \quad .$$

In der Regel werden die Randpunkte des Intervalls $[a, b]$ mit $a = u_0 < u_1 < \dots < u_j < \dots < u_{r+1} = b$ vorgeschrieben. Falls die zweiten Ableitungen für u_0 und u_{r+1} verschwinden, spricht man auch von natürlichen Spline-Funktionen.

Für die polynomiale L_2 -Approximation wurde im vorangehenden Abschnitt gezeigt, daß ein orthonormales Funktionensystem numerisch vorteilhaft ist. Eine analoge und nicht zuletzt auf praktische Überlegungen ausgerichtete Darstellung läßt sich auch für kubische Splines konstruieren.

Anstelle von (2.5.5.2-2) ist die Verwendung von ausgleichenden kubischen Splines in Form der "B-Splines" angezeigt ^(2/31).

$$(2.5.5.2-5) \quad c(t) = \sum_{\ell=-1}^{r+2} \lambda_{\ell} B_{\ell}(t) .$$

Die "B-Splines" haben die Differenzendarstellung

$$(2.5.5.2-6) \quad B_{\ell}(t) = \sum_{k=\ell-2}^{r+2} (t-u_k)_+^3 / \left[\prod_{\substack{q=\ell-2 \\ q \neq k}}^{\ell+2} (u_k - u_q) \right] .$$

Hierbei werden benachbarte Knoten durch

$$(2.5.5.2-7) \quad u_k = \begin{cases} u_1 - (1-k) \cdot (u_1 - t_{\min}) & \text{für } k \leq 0 \\ u_r + (k-r) \cdot (t_{\max} - u_r) & \text{für } k \geq m+1 \end{cases}$$

festgelegt. Die Schreibweise $(t-u_k)_+$ bedeutet wie üblich, daß der Ausdruck für $t \leq u_k$ verschwindet, andernfalls der Betrag einzusetzen ist.

Definitionsgemäß gilt für B-Splines

$$(2.5.5.2-8) \quad B_{\ell}(t) = 0 \quad \begin{cases} t > u_{\ell+2} \\ t < u_{\ell-2} \end{cases}$$

Sind die Knoten vorgegeben, führt dies auf ein lineares Approximationsproblem. Die Darstellung (2.5.5.2-5) - (2.5.5.2-8) hat den Vorteil, daß die Anzahl der unbekannten Parameter gleich der Anzahl der freien Parameter der Spline-Funktion ist. Weiterhin besitzt mit (2.5.5.2-8) die Koeffizienten-Matrix des Gleichungssystems heptagonale Gestalt, was bei großem r numerisch günstig ist.

Die Berechnung der Koeffizienten in (2.5.5.2-2) aus den B-Splines erfolgt rekursiv

$$(2.5.5.2-9) \quad c'''(u_i) = 6 d_i = \sum \lambda_{\ell} B_{\ell}'''(u_i) \quad (1 \leq i \leq r+1)$$

$$c''(u_i) = 2 c_i + 6 d_i u_i = \sum \lambda_{\ell} B_{\ell}''(u_i)$$

$$c'(u_i) = b_i + 2 c_i u_i + 3 d_i u_i^2 = \sum \lambda_{\ell} B_{\ell}'(u_i)$$

$$c(u_i) = a_i + b_i a_i + 2 c_i u_i^2 + d_i u_i^3 = \sum \lambda_{\ell} B_{\ell}(u_i)$$

Bei Festlegung von c in den Stützstellen und zwei weiteren Bedingungen ist das Gleichungssystem eindeutig lösbar.

Die Vorgabe der Randbedingungen und Stützstellen geschieht nach verschiedenen Kriterien. Schreibt man unterschiedliche Werte für die ersten und zweiten Ableitungen vor, lassen sich die resultierenden Effekte an den berechneten Kurvenverläufen deutlich erkennen (2/32).

Für praktische Anwendungen hat es sich als zweckmäßig erwiesen, möglichst die Zahl der Knoten gering zu halten und nicht mehr als einen Extremwert und einen Wendepunkt pro Segment zu haben. Nach Möglichkeit sollten Extrempunkte im mittleren Segmentbereich und Wendepunkte in Knotennähe positioniert sein (2/33).

Liegen (annähernd) äquidistante Knoten und viele Datenpunkte vor, dürfte es zweckmäßig sein, nach folgendem Vorgehen zu verfahren (2/34). Man definiert ein "Glättungsintegral"

$$(2.5.5.2-10) \quad \int_{u_0}^{u_{r+1}} c''(t)^2 dt$$

das unter der Bedingung

$$(2.5.5.2-11) \quad \sum_{i=0}^m \left[\frac{c(t_i) - s(t_i)}{\sigma(s(t_i))} \right] \leq k$$

zu minimieren ist. Durch Variation des Parameters k wird das Ausmaß der Glättung kontrolliert. Es wird vorgeschlagen, für $\sigma(s(t_i))$ eine Schätzung der Standardabweichung zu verwenden. In diesem Fall nimmt k einen innerhalb der Vertrauensbereiche liegenden Wert an

$$(m+1) - [2(m+1)]^{1/2} \leq k \leq (m+1) + [2(m+1)]^{1/2} \quad (2/35)$$

3 SIGNALDARSTELLUNG IM FREQUENZBEREICH

3.1 EINFÜHRUNG

Über das Studium psychophysiologischer Funktionen im Zeitbereich nehmen Signalanalysen im Frequenzbereich einen immer breiter werdenden Raum ein. Sie ermöglichen zusätzliche und unter Umständen entscheidende Einblicke in die Struktur und das Verhalten des zu untersuchenden psychophysiologischen Prozesses. Zeit- und Frequenzbereichsdarstellung sind eng miteinander verknüpft und ergänzen sich oft, wie dies beispielhaft aus der EEG- oder EKG-Diagnostik bekannt ist.

Obwohl eine Frequenzbereichsdarstellung nicht nur ausgehend von den trigonometrischen Funktionen (2.3.5.1) möglich ist, sondern ebenso mit Hilfe anderer orthogonaler und vollständiger Funktionensysteme erfolgen kann ^(3/1), wird für die nachfolgenden Punkte der Begriff der Fouriertransformation zugrundegelegt, da von hier aus die in der Praxis wichtigsten Prinzipien einer Frequenzbereichsdarstellung von Signalen beschrieben und unter Zuhilfenahme weiterer Information auch auf andere Funktionensysteme übertragen werden können. Sicherlich dürfte die Bedeutung der Fouriertransformierten für die Signalanalyse historisch begründbar sein, darüberhinaus jedoch besitzt sie eine Reihe von Eigenschaften, die für empirische und speziell psychophysiologische Daten ausgesprochen vorteilhaft sein können, wie im weiteren Verlauf zu zeigen ist ^(3/2).

3.2 FOURIERTRANSFORMATION

3.2.1 DEFINITION

Ist eine Signalfunktion $s(t)$ im Intervall $[-\infty, \infty]$ absolut integrierbar, d.h. gilt

$$(3.2.1-1) \quad \int_{-\infty}^{\infty} |s(t)| dt < \infty$$

und in $[-\infty, \infty]$ stückweise stetig ^(3/3), so bezeichnet man

$$(3.2.1-2) \quad s(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} s(\tau) \cos \omega(t-\tau) d\tau$$

auch als Fouriersche Integralformel (3/4). Setzt man nun

$$a(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} s(t) \cos \omega t \, dt \quad \text{und} \quad b(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} s(t) \sin \omega t \, dt,$$

ergibt sich für (3.2.1-2) der Ausdruck

$$(3.2.1-3) \quad s(t) = \int_0^{\infty} \{a(\omega) \cos \omega t + b(\omega) \sin \omega t\} d\omega$$

bzw.

$$(3.2.1-4) \quad s(t) = \int_0^{\infty} A(\omega) \cos(\omega t + \psi(\omega)) d\omega$$

oder

$$(3.2.1-5) \quad s(t) = \int_0^{\infty} A(\omega) \sin(\omega t + \phi(\omega)) d\omega$$

mit $A(\omega)$ definiert durch

$$(3.2.1-6) \quad A(\omega) = \{a^2(\omega) + b^2(\omega)\}^{1/2}, \quad \text{womit}$$

$$\cos \psi(\omega) = a(\omega)/A(\omega), \quad \cos \phi(\omega) = b(\omega)/A(\omega),$$

$$\sin \psi(\omega) = b(\omega)/A(\omega), \quad \sin \phi(\omega) = a(\omega)/A(\omega) \quad \text{sowie}$$

$$\psi(\omega) = \phi(\omega) + \frac{\pi}{2} \quad \text{sind.}$$

Addiert man desweiteren

$$0 = \frac{j}{2\pi} \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} s(\tau) \sin \omega(t-\tau) d\tau \quad \text{mit } j := \sqrt{-1}$$

zur Gleichung (3.2.1-2), in der lediglich die $\cos$ -Terme enthalten waren, ergibt sich unter Berücksichtigung der Eulerschen Beziehung

$$\cos \omega(t-\tau) + j \sin \omega(t-\tau) = e^{j\omega(t-\tau)}$$

der Ausdruck in komplexer Form

$$s(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} e^{j\omega(t-\tau)} s(\tau) d\tau \quad \text{bzw. umgeordnet}$$

$$(3.2.1-7) \quad s(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega t} d\omega \int_{-\infty}^{\infty} e^{-j\omega \tau} s(\tau) d\tau.$$

Wird zur Vereinfachung der Schreibweise berücksichtigt, daß $\omega = 2\pi f$ ist, und gilt

$$(3.2.1-8) \quad S(f) = \int_{-\infty}^{\infty} e^{-j2\pi f t} s(t) dt,$$

läßt sich (3.2.1-7) ausdrücken in der Form

$$(3.2.1-9) \quad s(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j2\pi f t} S(f) df.$$

Man bezeichnet nun mit (3.2.1-8) $S(f)$ allgemein als Fouriertransformierte der Funktion $s(t)$, während mit (3.2.1-9) $s(t)$ die in-

verse Fouriertransformierte von $S(f)$ heißt. In der Literatur findet sich anstelle von (3.2.1-8/9) auch die verkürzte Operator-schreibweise

$$S(f) = F \{s(t)\} \quad \text{und} \quad s(t) = F^{-1} \{S(f)\} \quad .$$

Gelten (3.2.1-8) und (3.2.1-9), bilden s und S ein sogenanntes Paar von Fouriertransformierten, was auch durch

$$s(t) \longleftrightarrow S(f)$$

symbolisiert wird.

Durch die bisher einleitende Darstellung der Fourierintegralformel und der Fouriertransformation dürften sich jetzt Parallelen zu der in (2.5.3.1) beschriebenen Reihe der trigonometrischen Funktionen und deren (physikalische) Interpretierbarkeit ziehen lassen.

Wird also angenommen, daß die Signalfunktion $s(t)$ die obige Voraussetzung (3.2.1-1) erfüllt, kann $s(t)$ unter Bezug auf (3.2.1-4) und (3.2.1-5) als Integral von harmonischen Vorgängen

$$\int A(\omega) \, d\omega \, \cos(\omega t + \psi(\omega))$$

$$\text{bzw.} \quad \int A(\omega) \, d\omega \, \sin(\omega t + \phi(\omega))$$

begriffen werden, wobei sich im betrachteten Intervall die Frequenzen fortlaufend ändern können. Damit sind die Ausdrücke in (3.2.1-6) anschaulich interpretierbar. Die Größe $A(\omega)$ meint die jeweilige Amplitude, während die Funktionen $\psi(\omega)$, $\phi(\omega)$ die ebenfalls von der Frequenz abhängigen Phasen bezeichnen.

Für numerische Zwecke wird insbesondere anstelle von (3.2.1-8/9) die komplexe Notation bevorzugt. Man zerlegt also $S(f)$ in einen Real- und Imaginärteil, d.h. es ist

$$(3.2.1-10) \quad S(f) = \operatorname{Re} S(f) + j \operatorname{Im} S(f) = |S(f)| e^{j\phi(f)}$$

und faßt dann den Betrag der Fouriertransformierten

$$(3.2.1-11) \quad |S(f)| = \{\operatorname{Re}^2 S(f) + \operatorname{Im}^2 S(f)\}^{1/2}$$

als das Spektrum der von f abhängigen Amplituden auf. Diese spektrale Interpretation der Fouriertransformierten wird noch in (3.3) konkreter beschrieben.

Durch die Berechnung von

$$(3.2.1-12) \quad \phi(f) = \arctan \frac{\operatorname{Im} S(f)}{\operatorname{Re} S(f)}$$

erhält man die zugehörigen Phasen.

3.2.2 EIGENSCHAFTEN

Die Fouriertransformierte besitzt u.a. die folgenden, nicht zuletzt für praktische Anwendungen wichtigen Eigenschaften.

(3.2.2-1) Linearität

Sind $S_1(f)$ und $S_2(f)$ die Fouriertransformierten der Funktionen $s_1(t)$, $s_2(t)$, so ergibt sich die Fouriertransformation für $(s_1(t) + s_2(t))$ als Summe der betreffenden einzelnen Fouriertransformierten, d.h.

$$\begin{aligned} \int_{-\infty}^{\infty} (s_1(t) + s_2(t)) e^{-j2\pi ft} dt &= \int_{-\infty}^{\infty} s_1(t) e^{-j2\pi ft} dt + \\ \int_{-\infty}^{\infty} s_2(t) e^{-j2\pi ft} dt &= S_1(f) + S_2(f) \end{aligned}$$

Außerdem gilt

$$\begin{aligned} F\{\alpha s(t)\} &= \alpha F\{s(t)\}, \text{ also} \\ \int_{-\infty}^{\infty} \alpha s(t) e^{-j2\pi ft} dt &= \alpha \int_{-\infty}^{\infty} s(t) e^{-j2\pi ft} dt = \alpha S(f) \end{aligned}$$

(3.2.2-2) Symmetrie

$s(t)$ und $S(f)$ bilden ein Paar von Fouriertransformierten, dann ist

$$S(t) \longleftrightarrow s(-f)$$

Aus Gleichung (3.2.1-9)

$$s(-t) = \int_{-\infty}^{\infty} S(f) e^{-j2\pi ft} df$$

und Vertauschung der Parameter t und f folgt

$$s(-f) = \int_{-\infty}^{\infty} S(t) e^{-j2\pi ft} dt$$

(3.2.2-3) Änderung der Einheit im Zeitbereich

Bilden $s(t)$ und $S(f)$ ein Paar, erhält man für $s(mt)$ mit $t_1 = mt$, wobei $m \in \mathbb{R}$, $m > 0$

$$\int_{-\infty}^{\infty} s(mt) e^{-j2\pi ft} dt = \int_{-\infty}^{\infty} s(t_1) e^{-j2\pi t_1 (f/m)} dt_1/m = \frac{1}{m} S\left(\frac{f}{m}\right)$$

Für $m < 0$ ändert sich das Vorzeichen im zweiten Term, insofern jetzt die Intervallgrenzen vertauscht sind. Somit ist

$$s(mt) \longleftrightarrow \frac{1}{|m|} S\left(\frac{f}{m}\right).$$

(3.2.2-4) Änderung der Einheit im Frequenzbereich

Die inverse Fouriertransformierte sei $S(mf)$ mit $m \in \mathbb{R}$. Eingesetzt in (3.2.1-9) erhält man mit $f_1 = mf$

$$\int_{-\infty}^{\infty} S(mf) e^{j2\pi ft} df = \int_{-\infty}^{\infty} S(f_1) e^{j2\pi f_1(t/m)} df_1/m \quad \text{und somit}$$

$$\text{ist} \quad \frac{1}{|m|} s\left(\frac{t}{m}\right).$$

(3.2.2-5) Verschiebung im Zeitbereich

$s(t)$ werde um den Betrag einer Konstanten t_1 verschoben. Mit $\tau = t - t_1$ berechnet sich

$$\begin{aligned} \int_{-\infty}^{\infty} s(t-t_1) e^{-j2\pi ft} dt &= \int_{-\infty}^{\infty} s(\tau) e^{-j2\pi f(\tau+t_1)} d\tau = \\ &= e^{-j2\pi ft_1} \int_{-\infty}^{\infty} s(\tau) e^{-j2\pi f\tau} d\tau = e^{-j2\pi ft_1} S(f) \end{aligned}$$

bzw. es ist in abgekürzter Schreibweise

$$s(t-t_1) \longleftrightarrow S(f) e^{-j2\pi ft_1}.$$

(3.2.2-6) Verschiebung im Frequenzbereich

Sei f_1 eine Konstante und $\xi = f - f_1$, dann ist mit (3.2.1-9)

$$\begin{aligned} \int_{-\infty}^{\infty} S(f-f_1) e^{j2\pi ft} df &= \int_{-\infty}^{\infty} S(\xi) e^{j2\pi(\xi+f_1)t} d\xi = \\ &= e^{j2\pi f_1 t} \int_{-\infty}^{\infty} S(\xi) e^{j2\pi \xi t} d\xi = e^{j2\pi f_1 t} s(t). \end{aligned}$$

Der Effekt einer solchen Verschiebung entspricht der Modulation von Signalen.

(3.2.2-7) Gerade und ungerade Fouriertransformierte

$s(t)$ wird als gerade Funktion bezeichnet, wenn $s(-t) = s(t)$ ist. Dies gilt beispielsweise für die cos-Funktion, d.h. es ist $\cos(t) = \cos(-t)$. Die zugehörige Fouriertransformierte ist ebenfalls eine gerade Funktion und sie ist reell

$$s(t) \longleftrightarrow \operatorname{Re} S(f) = \int_{-\infty}^{\infty} s(t) \cos(2\pi ft) dt.$$

Falls $s(-t) = -s(t)$ ist, wie dies für die sin-Funktion zutrifft, wird von einer ungeraden Funktion gesprochen ($\sin(t) = -\sin(-t)$).

Die Fouriertransformierte ist ungerade und imaginär

$$s(t) \leftrightarrow \text{Im } S(f) = -j \int_{-\infty}^{\infty} s(t) \sin(2\pi ft) dt$$

Die graphische Darstellung der Fouriertransformierten einer Signalfunktion getrennt nach Real- und Imaginärteil gibt m.a.W. bereits aufschlußreiche Hinweise.

3.2.3 BERECHNUNG

Da die Fouriertransformation sowohl für die Analyse psychophysiologischer Signale im Frequenzbereich als auch zur Vereinfachung von Berechnungen im Zeitbereich grundlegend ist, besitzt in der Praxis die Frage der effektiven numerischen Berechnung eine hervorragende Bedeutung.

Zur Reduzierung des numerischen Aufwandes bietet es sich zum einen an, die Symmetrien in den trigonometrischen Funktionen zu nutzen. Zum anderen ist eine geschickte Anordnung und Zusammenfassung von Teilsummen, die mit dem-selben Faktor multipliziert werden, angezeigt. Während klassische Ansätze vorwiegend von der Symmetrieeigenschaft Gebrauch machen, basieren auf dem letzteren Vorgehen die meisten der heute benutzten Methoden, die allgemein unter der Bezeichnung "schnelle Fouriertransformation" (Fast Fourier Transform = FFT) subsumiert werden. Das Grundprinzip der FFT ist insbesondere durch die Publikation von COOLEY & TUKEY im Jahre 1965 einem weiten Leser- und Benutzerkreis bekannt gemacht worden, läßt sich jedoch auf bereits 1905 von RUNGE geäußerte Überlegungen zurückführen (3/7). Daneben existieren Methoden, die beide Vorgehensweisen miteinander verbinden (3/8).

Durch Anwendung der FFT läßt sich die Zahl der sonst bei der direkten Berechnungen notwendigen m^2 komplexen Multiplikation (gemäß 3.2.1-8) auf $m \log_2 m$ bzw. $(m \log_2 m)/2$ komplexe Multiplikationen und $m \log_2 m$ komplexe Additionen reduzieren. Zur Verdeutlichung des grundlegenden Prinzips der FFT wird verkürzt der von COOLEY & TUKEY angegebene Algorithmus motiviert. Vorausgesetzt wird, daß die Meßwertanzahl m eine Potenz zur Basis 2 ist (3/9).

Anstelle der kontinuierlichen Version werde die Fouriertransformierte für den diskreten Fall betrachtet. Der Diskretisierung entspricht formal der Übergang von der bisherigen Integral- zur Summenschreibweise. Die diskrete Fouriertransformierte sei

$$(3.2.3-1) \quad S(v) = \sum_{t=0}^{m-1} s_0(t) e^{-j2\pi vt/m} \quad (0 < v < m-1) \quad .$$

Definiert man den Faktor a als

$$(3.2.3-2) \quad a = e^{-j2\pi/m}$$

erhält man anstelle von (3.2.3-1) den Ausdruck

$$(3.2.3-3) \quad S(v) = \sum_{t=0}^{m-1} s_0(t) a^{vt} \quad .$$

Unter der Bedingung, daß $m = 2^r$ ist, lassen sich v und t binär darstellen

$$(3.2.3-4) \quad \begin{aligned} v &= 2^{r-1} v_{r-1} + 2^{r-2} v_{r-2} + \dots + v_0 \quad , \\ t &= 2^{r-1} t_{r-1} + 2^{r-2} t_{r-2} + \dots + t_0 \quad . \end{aligned}$$

Sei weiterhin

$$(3.2.3-5) \quad w = (2^{r-1} v_{r-1} + 2^{r-2} v_{r-2} + \dots + v_0) (2^{r-1} t_{r-1} + 2^{r-2} t_{r-2} + \dots + t_0),$$

kann (3.2.3-3) zerlegt in Teilsummen wie folgt geschrieben werden:

$$(3.2.3-6) \quad S(v_{r-1}, v_{r-2}, \dots, v_0) = \sum_{t_0=0}^1 \sum_{t_1=0}^1 \dots \sum_{t_{r-1}=0}^1 s_0(t_{r-1}, \dots, t_0) a^w \quad .$$

Wegen

$$(3.2.3-7) \quad a^{\alpha+\beta} = a^\alpha a^\beta \quad (\alpha, \beta \in \mathbb{R})$$

berechnet sich

$$(3.2.3-8) \quad \begin{aligned} a^w &= a^{(2^{r-1} v_{r-1} + 2^{r-2} v_{r-2} + \dots + v_0) (2^{r-1} t_{r-1})} \\ &\quad a^{(2^{r-1} v_{r-1} + 2^{r-2} v_{r-2} + \dots + v_0) (2^{r-2} t_{r-2})} \\ &\quad a^{(2^{r-1} v_{r-1} + 2^{r-2} v_{r-2} + \dots + v_0) t_0} \quad . \end{aligned}$$

Aus der Definition (3.2.3-2) für a folgt nun

$$(3.2.3-9) \quad a^{2^r} = a^m = (e^{-j2\pi/m})^m = 1 \quad .$$

Damit berechnet sich der erste Term in (3.2.3-8) durch

$$(3.2.3-10) \quad a^{(2^{r-1}v_{r-1} + 2^{r-2}v_{r-2} + \dots + v_0)} (2^{r-1}t_{r-1}) = \\ \{a^{2^r(2^{r-2}v_{r-1}t_{r-1})}\} \{a^{2^r(2^{r-3}v_{r-2}t_{r-1})}\} \cdot \dots \\ \dots a^{2^r(v_1t_{r-1})} a^{2^{r-1}(v_0t_{r-1})} = a^{2^{r-1}(v_0t_{r-1})}$$

und entsprechend ergibt sich

$$(3.2.3-11) \quad a^{(2^{r-1}v_{r-1} + 2^{r-2}v_{r-2} + \dots + v_0)} (2^{r-2}t_{r-2}) = \\ \{a^{2^r(2^{r-3}v_{r-1}t_{r-2})}\} \{a^{2^r(2^{r-4}v_{r-2}t_{r-2})}\} \cdot \dots \\ \dots a^{2^{r-1}(v_1t_{r-2})} a^{2^{r-2}(v_0t_{r-2})} = a^{(2v_1+v_0)2^{r-2}t_{r-2}}.$$

Berücksichtigt man diesen Zusammenhang, erhält man anstelle von (3.2.3-6) den Ausdruck

$$(3.2.3-12) \quad S(v_{r-1}, v_{r-2}, \dots, v_0) = \sum_{t_0=0}^1 \sum_{t_1=0}^1 \dots \sum_{t_{r-1}=0}^1 s_0(t_{r-1}, t_{r-2}, \dots, t_0) \cdot \\ a^{2^{r-1}(v_0t_{r-1})} a^{(2v_1+v_0)2^{r-2}t_{r-2}} \dots \\ \dots a^{(2^{r-1}v_{r-1} + 2^{r-2}v_{r-2} + \dots + v_0)t_0}.$$

Betrachtet man nun die Einzelsummen, lautet (3.2.3-12) explizit (3.2.3-13)

$$s_1(v_0, t_{r-2}, \dots, t_0) = \sum_{t_{r-1}=0}^1 s_0(t_{r-1}, t_{r-2}, \dots, t_0) a^{2^{r-1}(v_0t_{r-1})} \\ s_2(v_0, v_1, t_{r-3}, \dots, t_0) = \sum_{t_{r-2}=0}^1 s_1(v_0, t_{r-2}, \dots, t_0) a^{(2v_1+v_0)2^{r-2}t_{r-2}} \\ \vdots \quad \quad \quad \vdots \\ s_r(v_0, v_1, \dots, v_{r-1}) = \sum_{t_0=0}^1 s_{r-1}(v_0, v_1, \dots, t_0) a^{2^{r-1}v_{r-1} + 2^{r-2}v_{r-2} + \dots + v_0)t_0}$$

$$S(v_{r-1}, v_{r-2}, \dots, v_0) = s_r(v_0, v_1, \dots, v_{r-1}).$$

Dieses System rekursiver Gleichungen (3.2.3-13) liegt dem von COOLEY & TUKEY angegebenen Algorithmus der schnellen Fouriertransformation zugrunde. Auf eine zusätzliche Auflistung des zugehörigen Signalflußgraphen wird ebenso wie auf die Angabe eines FFT-Programmes verzichtet, da an dieser Stelle das Grundprinzip der FFT erläutert werden sollte. Im übrigen dürften in jeder Programm-Bibliothek eines mittleren und größeren Rechenzentrums entsprechende FFT-Subroutinen implementiert sein, die aus Gründen der Effizienz unter Umständen eine maschinenabhängige Modifikation des oben skizzierten Prinzip einer möglichst günstigen Faktorisierung und Zusammenfassung darstellen (3/10).

3.3 CONVOLUTION UND SPEKTRUM

Als Faltung oder Convolution zweier Funktionen $s_1(t)$ und $s_2(t)$ bezeichnet man das Integral

$$(3.3-1) \quad s_1(\tau) * s_2(\tau) = \int_{-\infty}^{\infty} s_1(t) s_2(\tau-t) dt, \quad ,$$

wobei τ die Veränderliche und das Zeichen $*$ die Faltungsoperation symbolisiert (3/11).

Nach dem Faltungssatz ist die Fouriertransformierte $S_3(f)$ des Produktes zweier Funktionen $s_1(t), s_2(t)$ gleich der Convolution der zugehörigen Fouriertransformierten $S_1(f)$ und $S_2(f)$

$$(3.3-2) \quad S_3(f) = S_1(f) * S_2(f) = \int_{-\infty}^{\infty} S_1(\lambda) S_2(f-\lambda) d\lambda \quad .$$

Berücksichtigt man die Definitionen (3.2.1-8) und (3.2.1-9), kann man den Faltungssatz auch schreiben als

$$(3.3-3) \quad S_1(f) * S_2(f) = \int_{-\infty}^{\infty} S_1(\lambda) S_2(f-\lambda) d\lambda = \\ = \int_{-\infty}^{\infty} s_1(t) s_2(t) e^{-j2\pi ft} dt.$$

Für die inverse Fouriertransformation ergibt sich entsprechend

$$(3.3-4) \quad F^{-1}\{S_1(f) S_2(f)\} = \int_{-\infty}^{\infty} S_1(f) S_2(f) e^{j2\pi \tau f} df = \\ = \int_{-\infty}^{\infty} s_1(t) s_2(\tau-t) dt, \text{ m.a.W. ist (in verkürzter Schreib-}$$

$$\text{weise)} \quad F\{s_1(t) s_2(t)\} = F\{s_1(t)\} F\{s_2(t)\} \quad .$$

Aus (3.3-3) läßt sich das Parsevaltheorem für die Fouriertransformierte gewinnen

$$(3.3-5) \quad \int_{-\infty}^{\infty} s_1(t) s_2(t) dt = \int_{-\infty}^{\infty} S_1(f) S_2(-f) df$$

Wegen

$$S(-f) = S^*(f)$$

kann für den Fall $s_1(t) = s_2(t) = s(t)$ geschrieben werden

$$(3.3-6) \quad \int_{-\infty}^{\infty} s^2(t) dt = \int_{-\infty}^{\infty} S(f) S^*(f) df = \int_{-\infty}^{\infty} |S(f)|^2 df \quad \text{bzw.}$$

$$(3.3-7) \quad \int_{-\infty}^{\infty} s^2(t) dt = 2 \int_0^{\infty} |S(f)|^2 df$$

Bezogen auf die (physikalische) Interpretation dieser Ausdrücke stellt nun das Integral der rechten Seite als Integral über alle Frequenzen die Gesamtenergie einer Signalfunktion dar, was durch ein Beispiel kurz veranschaulicht werden soll.

Die Funktion $s(t)$ gebe den elektrischen Stromfluß an, der durch einen linearen (Ohmschen) Widerstand R fließt. Der Ausdruck $R \cdot s^2(t)$ bezeichnet dann die momentane Leistung (vgl. 2.3-14). Das Integral der linken Seite beziffert also bis auf den Faktor R die Gesamtenergie des Stroms $s(t)$. Dem Integral der rechten Seite entspricht die quadrierte Amplitudendichte pro Frequenz. Wiederum bis auf den Faktor R ist

$$(3.3-8) \quad E(f) = |S(f)|^2, \quad$$

d.h. die spektrale Energiedichte von $s(t)$.

Die Verwendung der Gleichungen (3.3-6/8) impliziert, daß $s(t)$ zeitlich begrenzt ist bzw. für $t \rightarrow \pm\infty$ hinreichend rasch gegen null abfällt. Vorausgesetzt wird in diesem Fall also, daß $s(t)$ eine Funktion der Art (2.2-1) bzw. (2.2-2) ist, deren Energie endlich und mittlere Leistung gleich null ist.

Eine gemäß (2.2-3) bestimmte Signalfunktion hingegen besitzt eine unendliche Gesamtenergie und eine nicht verschwindende mittlere Leistung. Die zugehörige spektrale Amplitudendichte im vorgegebenen Intervall $-P, P$ wird in Analogie zu (3.2.1-8) definiert durch

$$(3.3-9) \quad S(f, P) = \int_{-P}^P s(t) e^{-j2\pi ft} dt \quad \text{und mit}$$

$$(3.3-10) \quad E(f, P) = S(f, P) S^*(f, P) = |S(f, P)|^2$$

läßt sich die spektrale Energiedichte ermitteln.

Die Beträge von (3.3-9) und (3.3-10) wachsen monoton mit der Zunahme des Intervalls. Man betrachtet daher den Grenzwert und nennt

$$(3.3-11) \quad D(f) = \lim_{P \rightarrow \infty} E(f, P) / 2P$$

die spektrale Leistungsdichte. $D(f)$ ist also im Intervall $[-\infty, \infty]$ festgelegt.

Mit den obigen Beziehungen entspricht dann die Gesamtleistung D_{ges} dem Ausdruck

$$(3.3-12) \quad D_{\text{ges}} = \int_{-\infty}^{\infty} D(f) df .$$

Durch Grenzwertbildung erhält man speziell anstelle von (3.3-7) den Ausdruck

$$(3.3-13) \quad \lim_{P_1, P_2 \rightarrow \infty} \frac{1}{(P_1 - P_2)} \int_{-P_1}^{P_2} s^2(t) dt = \\ 2 \int_0^{\infty} |S(f)|^2 df = 2 \int_0^{\infty} D(f) df ,$$

der weiter unten im Text noch von Bedeutung ist.

3.4 SPEKTRALE LEISTUNGSDICHTE UND KORRELATIONSFUNKTION

3.4.1 AUTOKORRELATIONSFUNKTION UND SPEKTRALE LEISTUNGSDICHTE

Eine interessante und numerisch wichtige Beziehung zwischen der spektralen Leistungsdichte und der Korrelationsfunktion ergibt sich durch folgende Überlegung.

Die Autokorrelationsfunktion wurde bereits in Punkt 2.4 durch

$$(2.4-7) \quad \Omega_A(\tau) = \overline{s(t)s(t+\tau)} = \lim_{P \rightarrow \infty} \frac{1}{2P} \int_{-P}^P s(t)s(t+\tau) dt$$

definiert.

Für die in Punkt 3.3 notierte Deconvolution (3.3-4) erhält man unter der Bedingung, daß

$$S(-f) = S^*(f)$$

gilt, die Gleichung

$$(3.4.1-1) \quad \int_{-\infty}^{\infty} S_1(-f) S_2(f) e^{j2\pi f\tau} df = \int_{-\infty}^{\infty} s_1(t) s_2(t+\tau) dt .$$

Bei unperiodischen, stationären Signalfunktionen strebt $P \rightarrow \infty$, so daß man anstelle von (3.4.1-1) den Ausdruck

$$(3.4.1-2) \quad \lim_{P \rightarrow \infty} \int_{-\infty}^{\infty} S(-f, P) S(f, P) e^{j2\pi f\tau} df =$$

$$\lim_{P \rightarrow \infty} \frac{1}{2P} \int_{-P}^P s_1(t) s_2(t+\tau) dt$$

erhält. Setzt man nun $s_1(t) = s_2(t) = s(t)$, ergibt sich die Autokorrelationsfunktion aus

$$(3.4.1-3) \quad \Omega_A(\tau) = \lim_{P \rightarrow \infty} \frac{1}{2P} \int_{-P}^P s(t) s(t+\tau) dt =$$

$$\lim_{P \rightarrow \infty} \int_{-\infty}^{\infty} S(-f, P) S(f, P) e^{j2\pi f\tau} df .$$

Mit $S(-f) = S^*(f)$ und den Definitionen (3.3-10/11) für die spektrale Energie- und Leistungsdichte kann nun (3.4.1-3) auch geschrieben werden als

$$(3.4.1-4) \quad \Omega_A(\tau) = \lim_{P \rightarrow \infty} \frac{1}{2P} \int_{-P}^P s(t) s(t+\tau) dt =$$

$$\lim_{P \rightarrow \infty} \int_{-\infty}^{\infty} |S(f, P)|^2 e^{j2\pi f\tau} df = \int_{-\infty}^{\infty} D(f) e^{j2\pi f\tau} df ,$$

bzw. es gilt umgekehrt

$$(3.4.1-5) \quad D(f) = \int_{-\infty}^{\infty} \Omega_A(\tau) e^{-j2\pi f\tau} d\tau .$$

Das Leistungsspektrum einer stationären Signalfunktion ist m.a.W. gleich der Fouriertransformierten ihrer Autokorrelationsfunktion. Diese auch als Wiener-Khintchine Theorem bekannte Beziehung ist in der Praxis speziell bei umfangreichen Datensätzen von erheblichem numerischen Vorteil, wird ein effektiver Algorithmus zur Berechnung der Fouriertransformation vorausgesetzt ^(3/12).

3.4.2 KREUZKORRELATIONSFUNKTION UND SPEKTRALE KREUZLEISTUNGSDICHTE

Analog den Definitionen in (3.4.1) gibt die Fouriertransformierte der Kreuzkorrelationsfunktion (2.4-4)

$$(3.4.2-1) \quad D_{KR}(f) := \int_{-\infty}^{\infty} \Omega_{KR}(\tau) e^{-j2\pi f\tau} d\tau$$

die spektrale Kreuzleistungsdichte der betreffenden Signalfunktionen $s_1(t)$ und $s_2(t)$ an.

Durch Anwendung des Faltungssatzes aus (3.3) folgt speziell

$$(3.4.2-2) \quad \Omega_{KR}(\tau) = \lim_{P \rightarrow \infty} \frac{1}{P} \int_{-\infty}^{\infty} S_1(-f, P) S_2(f, P) e^{j2\pi f\tau} df,$$

wobei S_1 und S_2 die gemäß (3.3-9) definierten spektralen Leistungsdichten von $s_1(t)$ und $s_2(t)$ sind.

Die Umkehrung von (3.4.2-1) führt zu

$$(3.4.2-3) \quad \Omega_{KR}(\tau) = \int_{-\infty}^{\infty} D_{KR}(f) e^{j2\pi f\tau} df,$$

was wiederum analog der Schreibweise für die Autokorrelation (3.4.1-4) ist.

Zerlegt man die spektrale Kreuzleistungsdichte in einen Real- und Imaginärteil

$$(3.4.2-4) \quad D_{KR}(f) = C_{KR}(f) - j Q_{KR}(f) \quad (j = \sqrt{-1})$$

nennt man $C_{KR}(f)$ das Co-Spektrum und $Q_{KR}(f)$ das Quad-Spektrum.

Setzt man nun (3.4.2-4) in (3.4.2-3) ein und berücksichtigt, daß

$$D_{KR}(-f) = D_{KR}^*(f) = C_{KR}(f) + j Q_{KR}(f)$$

und $Q_{KR}(\tau)$ eine reellwertige Funktion ist, kann man die Kreuzkorrelationsfunktion auch schreiben in der Form

$$(3.4.2-5) \quad Q_{KR}(\tau) = \int_0^{\infty} [C_{KR}(f) \cos 2\pi f\tau + Q_{KR}(f) \sin 2\pi f\tau] df,$$

was die Bedeutung des Co- und Quad-Spektrums verständlicher werden läßt.

Bildet man für die Signalfunktionen $s_1(t, \Delta f)$ und $s_2(t, \Delta f)$, die im festen Frequenzintervall $[f_0, f_0 + \Delta f]$ definiert sind, die Kreuzkorrelationsfunktion, so kann man diese mit (3.4.2-5) schreiben als

$$(3.4.2-6) \quad \overline{s_1(t, f) s_2(t+\tau, \Delta f)} = 2\Delta f [C_{KR}(f) \cos 2\pi f\tau + Q_{KR}(f) \sin 2\pi f\tau].$$

Für den speziellen Fall, indem $\tau=0$ und $w \rightarrow 0$ ist, ergibt sich für das Co-Spektrum der Ausdruck

$$(3.4.2-7) \quad C_{KR}(f) = \lim_{f \rightarrow \infty} \frac{1}{2\Delta f} \overline{s_1(t, \Delta f) s_2(t, \Delta f)}$$

Führt man weiterhin eine Phasendrehung um 90° aus, was einem $\tau = \pi/2w$ entspricht, erhält man für das Quad-Spektrum

$$(3.4.2-8) \quad Q_{KR}(f) = \lim_{f \rightarrow \infty} \frac{1}{2\Delta f} \overline{s_1(t, \Delta f) s_2(t+\tau, \Delta f)}.$$

Durch diese Schreibweise wird deutlich, daß das Co-Spektrum die spektrale Leistungsdichte der phasengleichen Anteile in der Signalfunktion wiedergibt, während das Quad-Spektrum die Leistungsdichte der um 90° gegeneinander verschobenen Spektralanteile von $s_1(t)$ und $s_2(t)$ repräsentiert.

Die in den Kreuzleistungsspektren enthaltenen Phaseninformationen ermöglichen also eine aufschlußreiche Betrachtung und Analyse systemtheoretischer Beziehungen.

Für ein lineares System gelte die "Eingangs-Ausgangs-Relation"

$$(3.4.2-9) \quad s(t) = h(t) * z(t) = \int_{-\infty}^{\infty} h(u) z(t-u) du$$

mit $z(t)$ als Eingangs- und $s(t)$ als Ausgangssignal. Die Funktion $h(t)$ sei wiederum die Impulsantwort (Systemfunktion) und die Fouriertransformierte $H(f)$ von $h(t)$ die Übertragungsfunktion.

Korreliert man Eingangs- und Ausgangssignal gemäß (2.5-6)

$$\Omega_{KR}(\tau) = \overline{z(t) s(t+\tau)} = \lim_{P \rightarrow \infty} \frac{1}{2P} \int_{-P}^P z(t) s(t+\tau) dt$$

und setzt für $s(t)$ das obige Faltungsintegral (3.4.2-9) ein, ergibt sich

$$(3.4.2-10) \quad \Omega_{KR}(\tau) = \lim_{P \rightarrow \infty} \frac{1}{2P} \int_{-P}^P z(t) \int_{-\infty}^{\infty} h(u) z(t+\tau-u) du dt$$

bzw. nach Umordnung

$$(3.4.2-11) \quad \Omega_{KR}(\tau) = \int_{-\infty}^{\infty} h(u) \left[\lim_{P \rightarrow \infty} \frac{1}{2P} \int_{-\infty}^{\infty} z(t) z(t+\tau-u) dt \right] du .$$

Der Ausdruck in eckigen Klammern stellt die um den Betrag u verschobene Autokorrelationsfunktion (2.4-7) des Eingangssignals dar. Gleichung (3.4.2-11) läßt somit sofort erkennen, daß die Faltung der Autokorrelationsfunktion des Eingangssignals mit der Impulsantwort des Systems die Kreuzkorrelationsfunktion von Eingangs- und Ausgangssignal ergibt

$$(3.4.2-12) \quad \Omega_{KR}(\tau) = \int_{-\infty}^{\infty} h(u) \Omega_{AZ}(\tau-u) du = h(u) * \Omega_{AZ}(u)$$

bzw. nach Transformation in den Frequenzbereich

$$(3.4.2-13) \quad R_{KR}(f) = H(f) R_{AZ}(f)$$

mit $R_{KR}(f)$ und $R_{AZ}(f)$ als Fouriertransformierte der betreffenden Korrelationsfunktion.

Neben der Laufzeitanalyse durch Berechnung der Kreuzkorrelation bietet sich z.B. der in (3.4.2-12) notierte Zusammenhang zur Analyse systemtheoretischer Beziehungen an. So läßt sich etwa über die Kreuzkorrelationsfunktion die Impulsantwort (Systemfunktion) ermitteln, indem man als Eingangssignal ein weißes Rauschen wählt. Dessen spektrale Leistungsdichte besitzt voraussetzungsgemäß einen konstanten Wert, d.h. es ist

$$D_z(f) = D_0 = \text{const.} ,$$

und für die Autokorrelationsfunktion ergibt sich demnach

$$\Omega_{KR}(\tau) = D_0 \delta(\tau) .$$

Daraus folgt unmittelbar für die gesuchte Impulsantwort

$$h(\tau) = (1/D_0) \Omega_{KR}(\tau) .$$

In naheliegender Weise lassen sich die hier eingeführten Begriffe zur Untersuchung weiterer wichtiger systemtheoretischer Relationen auswerten.

Hierzu dient auch der Begriff der Kohärenz, der als normierte Kreuzleistungsdichte in der Praxis aus Gründen der besseren Vergleichbarkeit oft der Berechnung der normalen Kreuzleistungsdichte vorgezogen wird

$$(3.4.2-14) \quad \kappa(f) := \frac{D_{KR}(f)}{D_1(f) D_2(f)}$$

mit $D_1(f)$ und $D_2(f)$ als die gemäß (3.3-11) bestimmten Autoleistungsspektren der Funktionen $s_1(t)$, $s_2(t)$.

Analog der bisherigen Argumentation lassen sich die zwischen zwei Signalfunktionen gemeinsam auftretenden spektralen Anteile unter Bezug auf die Systemgleichung (3.4.2-9) analytisch nutzen.

3.5 LINEARE FILTERUNG

Die Filterung dient generell zur Trennung von Erwünschtem und Unerwünschtem, was je nach Anwendungsgebiet und -ziel sehr verschieden und in manchen Situationen nicht einmal hinreichend exakt definiert werden kann.

Im Rahmen der Analyse psychophysiologischer Signale werden Filter zur Glättung, Vorhersage, Kontrastverstärkung oder z.B. zur Identifizierung eines Signals aus einem Störpegel im Sinne eines Optimalfilters benutzt. Bekannt aus der routinemäßigen EEG- und EKG-Analyse dürfte vor allem die Kombination geeignet dimensionierter Hoch- und Tiefpassfilter sein, die es ermöglichen, aus einem komplexen Signal jenes Frequenzband herauszufiltern, das im wesentlichen nur noch die für die Fragestellung relevanten Signalanteile enthält.

In der Analogtechnik werden derartige Filter durch einfache RC-Schaltungen realisiert. Im allgemeinsten, abstrakten Fall lassen sich Filter mathematisch als Operatoren (Abbildungen) auf Funktionenräumen definieren. In dieser Arbeit wird die Betrachtung auf die praktisch wichtige Klasse der linearen Filter eingeschränkt, welche sich mit Hilfe der in (3.3) beschriebenen Faltungsoperation definieren lassen.

Die Faltung $s(\tau)$ zweier Funktionen z und h sei (3/13)

$$(3.5-1) \quad s(\tau) = z(\tau) * h(\tau) = \int_{-\infty}^{\infty} z(t) h(\tau-t) dt .$$

Die bei festem h hierdurch für jede Funktion z definierte Abbildung bezeichnet man auch als lineares System oder als linearen Filter und nennt z die Eingangsgröße (filter input), s die Ausgangsgröße (filter output) und die Gewichtsfunktion h Impulsantwort (filter response). Die Fouriertransformierte H der Gewichtsfunktion wird auch als Übertragungsfunktion (transfer function) angegeben.

Die Klasse der realisierbaren Filterfunktionen kann verständlicherweise sehr verschieden festgelegt werden. So wird die Klasse der Gewichtsfunktionen für Signale, die in Echtzeit zu verarbeiten sind, durch die als Folge des Kausalitätsprinzips erhobene Forderung der Realisierbarkeit eingeschränkt, nach der das Verhalten des Signals

nur von der Vergangenheit abhängig ist. Für $s(t)$ zum Zeitpunkt t_0 ist also nur $z(t)$ mit $t \leq t_0$ relevant. Es ist somit

$$(3.5-2) \quad h(t) = 0 \quad \text{für } t < 0.$$

Ist es hingegen möglich, die Signalfunktion zu speichern und dann erst zu verarbeiten, wie dies etwa bei einigen EKG-Analysatoren der Fall ist, entfällt diese Einschränkung.

So kann beispielsweise ein idealisierter Tiefpaßfilter gemäß (3.5-1) durch Wahl einer geeigneten Filterfunktion $h(t)$ bzw. deren Fouriertransformierten $H(f)$ verwirklicht werden. Für ein bandbegrenztes Signal verschwindet voraussetzungsgemäß die transformierte $S(f)$ unter- und oberhalb des Bandes. Sei f_0 die obere Grenzfrequenz, dann ist also

$$S(f) = 0 \quad \text{für } f > f_0.$$

Ein idealisierter Tiefpaßfilter hat demzufolge die Übertragungsfunktion

$$H(f) = \begin{cases} 1 & \text{für } f \leq f_0 \\ 0 & \text{sonst} \end{cases}.$$

Ein anderes Beispiel für einen linearen Filter stellt der sog. Comb-Filter dar. Als Beispiel sei auf ein periodisches Eingangssignal z mit der Periode P die Gewichtsfunktion

$$h(t) = \sum_{i=1}^m \delta(t - iP)$$

angewandt, wobei δ die Dirac'sche Deltafunktion bezeichnet (3/15).

Man erhält folglich als Ausgangssignal

$$s(\tau) = z * h = z(t) \sum_{i=1}^m \delta(\tau - t - iP) dt = \sum_{i=1}^m z(\tau - iP) = m z(\tau),$$

m.a.W. die Summe von m Perioden. Diese Filterung dient zur Verbesserung des Signal-zu-Rausch-Verhältnisses, was verständlich wird, wenn man das Signal als Repräsentanten eines stochastischen Prozesses auffaßt, der neben dem Signalanteil noch einen Rauschanteil enthält. Sind Signal und Rauschen unkorreliert, so läßt sich

mit Hilfe bekannter statistischer Verfahren unschwer zeigen, daß hierdurch das Signal-zu-Rausch-Verhältnis um den Faktor $\sqrt{m}$ verbessert wird. In der Praxis spielt dieser Filter insbesondere bei Untersuchungen über evozierte Potentiale eine besondere Rolle ^(3/16).

Formal entspricht der Filterung eines Signals auf einer digitalen Rechner also die Berechnung eines Faltungsintegrals im Zeitbereich, was jedoch auch auf einem leistungsfähigerem Computer noch relativ zeitaufwendig ist. Unter Bezug auf den Faltungssatz aus (3.3) ist es nun möglich, diese Operation durch einfache Multiplikation im Frequenzbereich durchzuführen, wenn - wie ebenfalls oben aufgezeigt - ein effektiver Algorithmus zur FFT verfügbar ist. In der Praxis wird man daher insbesondere bei vielen Daten die Filterung im Frequenzbereich bevorzugen. Abbildung 3.5-1 gibt die beiden Vorgehensweisen der Realisierung eines linearen Filters auf einer digitalen Anlage schematisch wieder ^(3/17).

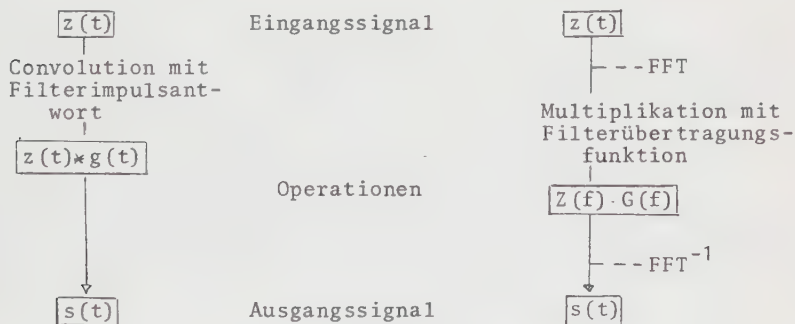


Abb. 3.5-1 LINEARE FILTERUNG

Die Berechnung eines Filters wirft insofern ein bisher noch nicht thematisiertes Problem auf, da nicht kontinuierliche Signalfunktionen, sondern Funktionen von endlich vielen diskreten Werten betrachtet und numerisch ausgewertet werden, worauf in dem jetzt anschließenden Punkt 3.6 kurz eingegangen werden soll.

3.6 ASPEKTE DER NUMERISCHEN AUSWERTUNG

3.6.1 ABTASTUNG UND SEGMENTIERUNG

Die numerische Auswertung von Signalfunktionen auf einer digitalen Rechenanlage impliziert, daß es sich bei den zu analysierenden Daten um diskrete, endlich viele Funktionswerte endlicher Zifferzahl handelt ^(3/18). Praktisch ist die an diesem Punkt der Datenauswertung erforderliche Diskretisierung kontinuierlicher Signalfunktionen nicht unwichtig, als der kontinuierlichen Funktion eine Mindestanzahl von Meßpunkten entnommen werden müssen, soll die im aufgenommenen Signal enthaltene Information wiedergegeben werden. Nach dem Abtast- oder Nyquist-Theorem ^(3/19) muß die Abtastfrequenz f_T mindestens doppelt so hoch sein wie die höchste im Signal enthaltene, untersuchungsrelevante Frequenz f_U

$$(3.6.1-1) \quad f_T > 2 \cdot f_U$$

Damit ergibt sich für das Signal, welches die gewünschte Information enthält, der Ausdruck

$$s(t) = \sum_i s(i/2f_U) \sin(2f_U t - i) / [(2f_U i - k)\pi]$$

und für die Anzahl m der im betrachteten Intervall der Länge T abzutastenden Punkte

$$m > 2 f_U T .$$

Abgesehen davon, daß die Abtastung nicht instantan erfolgt und nur endlich viele Ziffern im Computer darstellbar sind, ergeben sich Probleme bzw. Abweichungen in der Praxis dadurch, daß mitunter die Zahl der erforderlichen Meßpunkte kleiner ist, als durch f_U vorgegeben wird. In diesem Fall erhält man also lediglich für Frequenzen $< f_U$ zuverlässige Informationen.

Desweiteren kann eine Verfälschung der Signalfunktion dadurch hervorgerufen werden, daß man aus technischen Gründen das Gesamtintervall in bestimmte Teilintervalle aufzuteilen gezwungen ist. Die resultierenden Artefakte lassen sich durch die Wahl einer geeigneten

Fensterfunktion zumindest teilweise wieder aufheben, wobei in der Regel ein Kompromiß zwischen der maximal erreichbaren Frequenzauflösung und der nicht vermeidbaren Verzerrung zu schließen ist.

Zur Verringerung der durch eine Segmentierung des aufgenommenen Signals auftretenden Unterschiede von berechnetem und wahren Spektrum, welche praktisch durch eine Faltung mit dem Spektrum der Segment- oder auch Ausblendfunktion bedingt ist, lassen sich beispielsweise die folgenden Gewichtsfunktionen verwenden, mit denen das Eingangssignal multipliziert wird und die als Fensterfunktionen oder ebenso als spezielle Klasse von Filtern gemäß der Beschreibung in (3.5) bezeichnet werden

$$(3.6.1-2) \quad h(t) = 1 - |2 \cdot t| \quad (\text{Dreiecksfenster})$$

$$(3.6.1-3) \quad h(t) = 1/2 [1 - \cos(2\pi t)] \quad (\text{Hanning-Fenster})$$

$$(3.6.1-4) \quad h(t) = e^{-t^2} \quad (t \in [-\frac{1}{2}, \frac{1}{2}]) \quad .$$

Die Auswahl einer angemessenen Filterfunktion nach präzisierten Kriterien und die sich ergebenden Effekte auf die Gestalt der Spektren sollen hier nicht näher behandelt werden, da sich ausführlichere Diskussionen an zahlreichen Stellen der angegebenen Literatur finden (3/20), wohingegen abschließend die für eine Signalanalyse im Zeit- und Frequenzbereich wesentlichen, für praktische Berechnungen wichtige Schritte in schematischer Form zusammengestellt werden.

3.6.2 SCHRITTE DER BERECHNUNG

In den vorangegangenen Punkten ist auf eine Reihe von Beziehungen zwischen der Faltung zweier Funktionen, der Auto- und Kreuzkorrelationsfunktion und den Fouriertransformierten der entsprechenden involvierten Funktionen hingewiesen worden. Die dort beschriebenen Zusammenhänge lassen sich für die numerische Auswertung auf digitalen Rechenanlagen vorteilhaft nutzen, wird ein effektiver Algorithmus zur FFT - wie in 3.2.3 aufgezeigt - für die Berechnungen zugrundegelegt.

So wird man die Convolution zweier Funktionen z und h gemäß den in Abb. 3.6.2-1 angeführten Schritten berechnen.

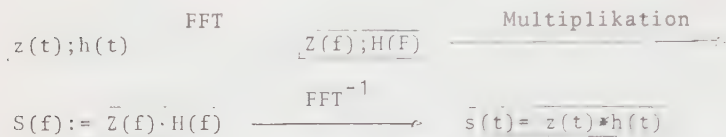


Abb. 3.6.2-1 SCHRITTE ZUR BERECHNUNG DER CONVOLUTION ZWEIER FUNKTIONEN

Würde man das Faltungsintegral $z * h$ bzw. die durch Diskretisierung hieraus entstandene Summe für n Meßpunkte berechnen, wären ca. n^2 Additionen und Multiplikationen notwendig. Da - wie auch in (3.2.3) erläutert - der Aufwand für eine FFT im günstigen Fall bei $n \cdot \log_2(n)$ liegt, ergibt obiges Rechenschema also ($3 \times \text{FFT} + n$ Multiplikation $=$) $3 n \cdot \log_2(n) + n$ Operationen. Für $n = 1024 = 2^{10}$ entsteht m.a.W. im ersten Fall ein Rechenaufwand von ca. 10^6 Punktoperationen, während im zweiten Fall nur etwa $3,1 \cdot 10^4$ Punktoperationen notwendig sind, was eine Rechenzeitreduzierung auf ungefähr $1/30$ bedeutet. Darüberhinaus verringern sich die auftretenden Rundungsfehler erheblich (3/21).

Die Autokorrelation (3.4.1-5) und Kreuzkorrelation (3.4.2-2) einer Signalfunktion wird man entsprechend den unten skizzierten Schritten berechnen.

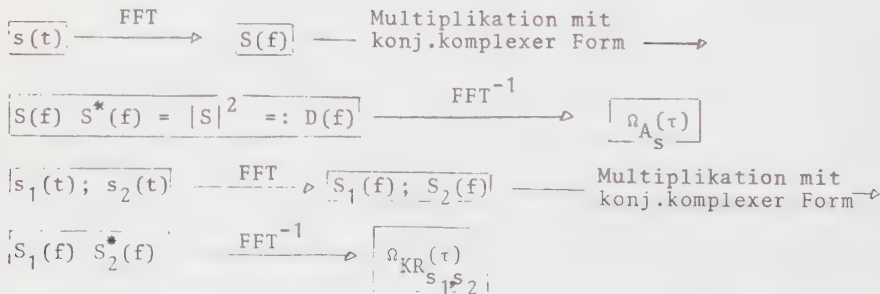


Abb. 3.6.2-2 SCHRITTE ZUR BERECHNUNG DER AUTO- UND KREUZKORRELATIONSFUNKTION

4 NUMERISCHE KLASSIFIKATION

4.1 EINFÜHRUNG

4.1.1 PROBLEMSTELLUNG

Die Entwicklung effektiver Algorithmen wie etwa der schnellen Fourier-, Walsh-Transformation oder der schnellen Convolution sowie die Installierung leistungsfähiger Kleinrechnersysteme führen im wissenschaftlichen und klinischen Bereich zu einer verstärkten Aufnahme, Beschreibung und Analyse von Biosignalen. Um die zahlenmäßig weit angestiegenen Daten nicht lediglich archivieren zu müssen, sondern um die sich in dieser Situation eröffnenden Möglichkeiten zur Verbesserung der diagnostischen, therapeutischen und prognostischen Aussagen angemessen nutzen zu können, werden über die bisher behandelten Methoden der Signalbeschreibung im Zeit- und Frequenzbereich hinaus zunehmend mathematisch bzw. statistisch anspruchsvollere Verfahren und Modelle erforderlich.

Berücksichtigt man speziell die durch den Fortschritt psychophysiologischer Projekte aufgeworfenen Fragestellungen der differenzierten Bestimmung des Effekt-Nachweises, der Effekt-Determination und der Effekt-Generalisation ^(4/1), dürfte man leicht jene Bedeutung abschätzen können, welche in einem immer größeren Ausmaß multivariate Strategien der Datenreduktion und Datenklassifikation erlangen.

Im Unterschied zur Beschreibung einer einzelnen oder weniger Biosignalfunktionen werden bei multivariaten Ansätzen mehrere Signalfunktionen betrachtet. Die meist umfangreiche Menge der untersuchten Signalfunktionen soll hierbei sowohl intraindividuell als auch interindividuell eine hinreichend repräsentative Stichprobe der relevanten Population darstellen ^(4/2).

So ist in der Psychophysiologie z.B. nicht nur die im Einzelfall erfolgreich durchgeführte Symptomkontrolle bei psychosomatischem Herzbeschwerden durch Methoden des Biofeedback von Interesse, sondern steigende Aufmerksamkeit wird umfassenderen wissenschaftlichen Studien entgegengebracht, in denen an einem

repräsentativen Untersuchungskollektiv die therapeutische Relevanz lerntheoretisch ausgerichteter Interventionen ermittelt werden kann.

Eine der in derartigen größeren Forschungsvorhaben zentralen Schwierigkeiten liegt oft darin begründet, daß eine übersichtliche und d.h. interpretierbare Darstellung der in komplexer Form vorliegenden Information über die jeweiligen psychophysiologischen Funktionsverläufe nur schwer zu finden ist.

Das methodische und methodologische Problem einer adäquaten Datenbeschreibung und -analyse kann man zum einen zu lösen versuchen, indem man die Zahl der erhobenen Messwerte oder Merkmale reduziert. Dieses Ziel verfolgen Verfahren der Parameterextraktion bzw. Verfahren der Dimensionsstatistik (4/5).

Zum anderen wird man eine zweckmäßige Aufteilung der Signalfunktionen in bestimmte Klassen oder Gruppen anstreben, um so einen Einblick in die den Daten inhärente Struktur zu erhalten.

Da in der Psychologie Verfahren der Dimensionsstatistik relativ häufig beschrieben und benutzt werden, soll im folgenden das Problem einer optimalen Zerlegung einer vorgegebenen Elementenmenge in Teilmengen kurz aufgegriffen werden. Eine stärkere Verwendung der diesem Kontext zuzurechnenden Verfahren dürfte nicht zuletzt für die Psychophysiologie zu aufschlußreichen Resultaten beitragen.

4.1.2 PRÄZISIERUNG DER PROBLEMSTELLUNG

Während in Situationen, in denen eine Aufteilung der gegebenen Menge von Biosignalfunktionen bereits existiert, lediglich einen effektiven Algorithmus zur möglichst sicheren Zuordnung der einzelnen Funktionen zu den vorgegebenen Klassen zu entwickeln ist, besteht bei der numerischen Klassifikation das eigentliche Problem darin, eine solche Klasseneinteilung erst zu konstruieren.

In Erweiterung der diskriminanzanalytischen Fragestellung handelt es sich bei dem Klassifikationsproblem also darum, eine vorgegebene Menge von Elementen in möglichst gut separierte Teilmengen, Klassen oder Gruppen zu zerlegen. Die zugehörigen Ansätze, Verfahren und Algorithmen werden als Klassifikationsverfahren bezeichnet. In neuerer Zeit wird diese Gruppe mathematischer, statistischer und heuristischer Verfahren auch unter dem allgemeineren Begriff "Clusteranalyse" subsumiert (4/4).

Zur Vereinfachung der nachfolgenden Übersicht über Ansätze und Verfahren der numerischen Klassifikation oder auch Clusteranalyse wird die bisher betrachtete Biosignalfunktion $s(t)$ als Zufallsvektor x' aufgefaßt, dessen m -Komponenten die in den Punkten $t_1, t_2, \dots, t_m$ registrierten Meßwerte repräsentieren. Durch Normierung gelte zudem

$$m_i = m_j = m$$

Die Indices i, j bezeichnen die Vektoren i, j , d.h. - wenn nicht anders angegeben - besitzen die Signalfunktionen $s_i(t)$ und $s_j(t)$ gleiche Meßwertanzahl, was im übrigen leicht durch geeignete Transformation zu erreichen ist.

Die in Frage kommenden n -Biosignalfunktionen lassen sich in einer Matrix

$$(4.1.2-1) \quad X = (x_{il}) = \begin{pmatrix} x_{11} & \dots & x_{1l} & \dots & x_{1m} \\ \vdots & & \vdots & & \vdots \\ x_{il} & \dots & x_{il} & \dots & x_{im} \\ \vdots & & \vdots & & \vdots \\ x_{n1} & \dots & x_{n1} & \dots & x_{nm} \end{pmatrix}$$

zusammenfassen.

Für die folgende verkürzte Darstellung werden also stochastische Biosignale der Form (2.2.7) zugrundegelegt, für die lediglich eine andere Schreibweise gewählt wird.

Die Menge der Objekte (Personen), an denen die Funktionen gemessen worden sind, wird einheitlich mit E angegeben,

$$(4.1.2-2) \quad E = \{e_i \mid 1 \leq i \leq n\}$$

Das Element e_i stellt also die i -te Person dar, dessen Meßwerte in dem m -dimensionalen Vektor $x_i^!$ wiedergegeben werden.

Die Aufgabe, eine optimale Zerlegung von E in Teilmengen zu konstruieren, könnte man dem intuitiven Vorgehen folgend dadurch zu lösen versuchen, daß man die verschiedenen Zerlegungen von E erstellt und diese nach einem geeigneten Kriterium miteinander vergleicht. Inwieweit eine solche Strategie realistisch ist, erkennt man unmittelbar anhand der Zahl der für n Elemente möglichen Zerlegungen von E .

Sei E die nichtleere, endliche Menge (4.1.2-2) und G eine Zerlegung (Partition) von E in $k \leq n$ elementfremde Teilmengen g_i , deren Vereinigung E ergibt

$$(4.1.2-3) \quad E = \bigcup_{i=1}^k g_i \quad \text{und} \quad g_i \cap g_j = \emptyset \quad (i, j=1, \dots, k; i \neq j).$$

Dann berechnet sich die Anzahl $S(n, k)$ der bei k Teilmengen möglichen Partitionen G von E über die Formel

$$(4.1.2-4) \quad S(n, k) = \frac{1}{k!} \cdot \sum_{i=0}^k (-1)^i \cdot \binom{k}{i} \cdot (k-i)^n$$

bzw. im Anwendungsfall einfacher rekursiv über

$$S(n+1, k) = S(n, k-1) + k \cdot S(n, k) \quad \text{mit} \quad S(n, 1) = S(n, n) = 1 \\ S(n, k) = 0 \quad \text{für} \quad n < k.$$

Für den Spezialfall $k=2$ ergibt sich also

$$S(n, 2) = 2^{n-1} - 1 \quad (4/5).$$

Die $S(n, k)$ werden auch Stirling'sche Zahlen zweiter Art genannt und besitzen etwa für $n = \{1, 2, \dots, 10, 15, 20, 25, 50, 100\}$ und $k = \{2, 3, 4, 5, 10\}$ folgende Werte:

N	K \ 1	2	3	4	5	10
1	1	0	0	0	0	0
2	1	1	0	0	0	0
3	1	3	1	0	0	0
4	1	7	6	1	0	0
5	1	15	25	10	1	0
6	1	31	90	65	15	0
7	1	63	301	350	140	0
8	1	127	966	1701	1050	0
9	1	255	3025	7770	6951	0
10	1	511	9330	34105	42525	1
15	1	16393	2375101	42355950	2108 E09	12662650
20	1	524297	560606446	4523 E07	7492 E08	5918 E09
25	1	1678 E04	1412 E08	4677 E10	2437 E12	1203 E15
50	1	5629 E11	1196 E22	5282 E25	7401 E29	2615 E40
100	1	6338 E23	9590 E43	6704 E55	6574 E64	2756 E90

YY

ANMERKUNG: XXXX EYY BEDEUTET :XXXX*10

Abb. 4.1.2.1 Stirling'sche Zahlen zweiter Art

Wird die Anzahl k der Gruppen nicht vorherbestimmt, kann man die Anzahl B_n aller möglichen Partitionen rekursiv ermitteln, indem man folgende Formel benutzt

$$B_n = \sum_{i=1}^n S(n, i),$$

woraus sich für $n \geq 0$ ergibt:

$$B_{n+1} = \sum_{i=0}^n \binom{n}{i} \cdot B_i \quad \text{mit} \quad B_0 := 1; B_1 = 1.$$

Diese Zahlen werden Bell'sche Zahlen oder Exponentialzahlen genannt.

N	B(N)	N	B(N)
1	1	15	1383 E06
2	2	20	5172 E10
3	5	25	4639 E14
4	15	30	8467 E20
5	52	35	2816 E26
6	203	40	1575 E32
7	877	45	1393 E38
8	4140	50	1857 E44
9	21147	60	9769 E56
10	115975	70	1807 E70

Abb. 4.1.2.2 Bell'sche Zahlen

Man erkennt, daß schon für kleine Werte von n und k die Anzahl der Gruppierungen ins enorme wächst, so daß die Methode der totalen Enumeration bereits für kleine Fallzahlen selbst angesichts der auf einer elektronischen Rechanlage realisierbaren Möglichkeiten zu aufwendig ist. So ergibt sich die Notwendigkeit zur Entwicklung effizienter Klassifikationsverfahren.

Orientiert an Methoden der Optimierung wie der dynamischen Programmierung oder der Branch and Bound-Verfahren ^(4/6) lassen sich Strategien formulieren, durch die eine hohe Zahl überflüssiger Partitionen eliminiert und das Optimum aus der restlichen, geringeren Zahl erreicht werden kann. Im Unterschied zu diesen bisher nur für sehr kleine Fallzahlen praktikablen Ansätzen führen die meisten Klassifikationsverfahren von einer plausibel erscheinenden Optimierungsstrategie ausgehend nur zu lokalen (suboptimalen) Lösungen. Diese heuristischen Verfahren werden dennoch oft benutzt, da das globale und eventuell nur um geringe Prozent bessere globale Optimum nur durch einen weitaus höheren Rechenaufwand zu erzielen ist.

4.2 DISTANZFUNKTIONEN

4.2.1 DEFINITION

Die Konstruktion einer optimalen Partition G von E setzt die Vergleichbarkeit der Elemente voraus ^(4/7). Eine Distanz d kann als reellwertige Funktion $d: E \times E \rightarrow \mathbb{R}_+$ aufgefaßt werden, durch die den Elementen $e_i, e_j \in E$ eine nicht-negative reelle Zahl $d = d(e_i, e_j) = d_{ij}$ zugeordnet wird. Es gilt

$$(4.2.1-1) \quad d_{ij} \geq 0$$

$$(4.2.1-2) \quad d_{ij} = d_{ji}$$

$$(4.2.1-3) \quad d_{ii} = 0.$$

Sind zusätzlich die Forderungen

$$(4.2.1-4) \quad d_{ij} = 0 \rightarrow e_i = e_j$$

$$(4.2.1-5) \quad d_{ie} \leq d_{ij} + d_{je}$$

erfüllt, liegt eine metrische Distanzfunktion oder kurz Metrik vor (4/8).

Die $(n \times n)$ Matrix $D = (d_{ij})$ heißt Distanzmatrix. Sie ist wegen (4.2.1-2) symmetrisch und wegen (4.2.1-3) sind ihre Diagonalelemente $d_{ii} = 0$.

Die Definition der Ähnlichkeit (Similarity) zwischen e_i, e_j erfolgt analog durch $s: E \times E \rightarrow R$ mit $s = s(e_i, e_j) = s_{ij}$ (4/9).

Zur Transformation von Distanzmaßen in Ähnlichkeitsmaße bzw. umgekehrt werden u.a. benutzt

$$(4.2.1-6) \quad s = 1/(1+d^2) \quad \text{mit } a = \max d_{ij} \quad (i \neq j)$$

$$(4.2.1-7) \quad s = 1-d/a$$

$$(4.2.1-8) \quad s = (H-d^2)/(h+d^2) \quad \text{mit } H = \text{mittlere Distanz.}$$

Speziell für $(0 \leq s \leq 1)$ gilt

$$(4.2.1-9) \quad d = 1-s$$

$$(4.2.1-10) \quad d = (1-s)^{\frac{1}{2}}$$

$$(4.2.1-11) \quad d = -\log s .$$

Nicht selten wird für $-1 \leq s \leq +1$ die Transformation

$$(4.2.1-12) \quad d = (2(1-s))^{\frac{1}{2}}$$

benutzt (4/10).

Auf naheliegende Weise läßt sich eine auf $E \times E$ definierte Distanzfunktion zu einer Distanzfunktion $d: E^* \times E^* \rightarrow R_+$ (mit E^* als Potenzmenge von E) erweitern, welche die Distanzen zwischen den Teilmengen von E angibt.

4.2.2 NICHT-METRISCHE DISTANZFUNKTIONEN

In der Tabelle 4.2.2.1 sind die Kombinationsmöglichkeiten für binäre Variablen aufgeführt.

		e_i		
		1	0	
e_j	1	a	c	a+c
	0	b	e	b+e
		a+b	c+e	m=a+b+c+e

Tab. 4.2.2.1 (2,2)-Tabelle der Kombinationsmöglichkeiten binärer Variablen

Die hier vorgeschlagenen Ähnlichkeitsfunktionen ergeben sich teilweise als Spezialfall der Funktion

$$(4.2.2-1) \quad s_{ij}^{\delta\lambda} = \frac{a+\delta e}{a+\delta e+\lambda(b+c)} \quad .$$

Die Konstanten $0 \leq \delta \leq 1$ $\lambda > 0$ stellen geeignet zu wählende Gewichtungsfaktoren dar. So bedeutet $\delta=0$ etwa, daß gemeinsames Fehlen von Eigenschaften nicht, und $\delta=1$, daß gemeinsames Fehlen von Eigenschaften voll berücksichtigt wird. Die Zahl 2 hebt je nach Größe das Übereinstimmen ($\lambda < 1$) oder aber das Nichtübereinstimmen ($\lambda > 1$) hervor.

Die Funktion (4.2.2-1) läßt sich auf nominale Variablen mit $r > 2$ Kategorien oder Zuständen erweitern. Analog der Definition der Größen a,b,c,e in Tab. 4.2.2.1 bezeichnet

- m_{a+e} Anzahl der Variablen, bei denen die Elemente übereinstimmen.
- m_e Anzahl der Variablen, bei denen die Elemente in einer unerwünschten Kategorie (etwa im Fehlen von Eigenschaften) übereinstimmen.
- m_{b+c} Anzahl der Variablen, bei denen die Elemente nicht übereinstimmen.

Eine ganze Schar von Ähnlichkeitsfunktionen ergibt sich durch

$$(4.2.2-2) \quad s_{ij}^{\delta\lambda} = \frac{m_{a+e}^{-(1-\delta)} m_e}{m_{a+e}^{-(1-\delta)} m_e + \lambda m_{b+c}}$$

Ähnlich zahlreich sind die bei ordinalen Variablen möglichen Definitionen einer geeignet erscheinenden Ähnlichkeit bzw. Distanz. Seien für die Variablen außer $x_i = x_j$, $x_i \neq x_j$ noch die Beziehungen $x_i < x_j$ und $x_i > x_j$, daher $\min(x_i, x_j)$ oder $\max(x_i, x_j)$ sinnvoll, lassen sich etwa Distanzfunktionen der Form

$$(4.2.2-3) \quad d_{ij} := \frac{\sum_{l=1}^m \frac{|x_{il} - x_{jl}|}{(x_{il} + x_{jl})}}$$

bzw.

$$(4.2.2-4) \quad d_{ij} := \frac{\sum_{l=1}^m \frac{|x_{il} - x_{jl}|}{\sum_{l=1}^m (x_{il} + x_{jl})}}$$

definieren (4/11). Neben probabilistischen Ähnlichkeitsmaßen wird u.a. die metrische Distanzfunktion

$$(4.2.2-5) \quad d_{ij} := \frac{\sum_{l=1}^m x_{il} + \sum_{l=1}^m x_{jl} - 2 \sum_{l=1}^m \min(x_{il}, x_{jl})}{\sum_{l=1}^m x_{il} + \sum_{l=1}^m x_{jl} - \sum_{l=1}^m \min(x_{il}, x_{jl})}$$

benutzt (4/12).

Abhängig vom Untersuchungsziel wird man eine der hier erwähnten bzw. weitere und unschwer zu formulierende Distanz- oder Ähnlichkeitsfunktionen anwenden (4/13).

4.2.3 METRISCHE DISTANZFUNKTIONEN

Bei quantitativen Merkmalen wird man die Elemente e_i, e_j durch ihre Meßwertvektoren

$$x_i' = (x_{i1}, x_{i2}, \dots, x_{im})$$

$$x_j' = (x_{j1}, x_{j2}, \dots, x_{jm})$$

als Punkte im m -dimensionalen reellen Zahlenraum R^m auffassen.

Meist wird für quantitative Variablen der euklidische Abstand

$$(4.2.3-1) \quad d_{ij} := \|x_i - x_j\|_2 = \left| (x_i - x_j)' (x_i - x_j) \right|^{\frac{1}{2}} = \left| \sum_{l=1}^m (x_{il} - x_{jl})^2 \right|^{\frac{1}{2}}$$

verwandt. Diese metrische Distanz ist translationsinvariant und invariant unter orthogonalen linearen Transformationen der Meßwertvektoren, da

$$\|y_i - y_j\|_2 = \|x_i - x_j\|_2$$

gilt, falls $y_i = C x_i + b$ und $y_j = C x_j + b$ mit $b \in R^m$ und C eine orthogonale $(m \times m)$ -Matrix ist, wie man wegen $C'C = I$ leicht verifiziert.

(4.2.3-1) ist allerdings nicht skaleninvariant und nicht invariant unter beliebigen nichtsingulären Transformationen. Man wird daher die Einheiten der m Meßwerte vergleichbar wählen, was unter Umständen durch geeignete Transformation

$$(4.2.3-2) \quad y_{il} := \frac{(x_{il} - \min_i x_{il})}{(\max_i x_{il} - \min_i x_{il})}$$

($i = 1, \dots, n; l = 1, \dots, m$)

erreicht wird. Des öfteren wird hierzu auch die Normierung

$$(4.2.3-3) \quad z_{i1} := (x_{i1} - x_{.1}) / s_1$$

$$\text{mit } x_{.1} := \frac{1}{n} \sum_{i=1}^n x_{i1} \quad \text{und } s_1^2 := \frac{1}{n} \sum_{i=1}^n (x_{i1} - x_{.1})^2$$

verwandt, welche allgemein als z-Transformation bekannt ist.

Da stark korrelierende Variablen zu erheblichen Verzerrungen der angestrebten Klassifikation führen können, bietet sich anstelle der euklidischen Distanz die MAHALANOBIS-Distanz an.

Die (empirische) Kovarianzmatrix sei

$$K := \frac{1}{n} \sum_{l=1}^n (x_l - \bar{x})(x_l - \bar{x})'$$

mit

$$k_{ij} = \frac{1}{n} \sum_{l=1}^n (x_{li} - \bar{x}_{.i})(x_{lj} - \bar{x}_{.j})$$

Die Existenz der Inversen von K werde vorausgesetzt (4/14). Mit $\lambda_i > 0$ d.h. der Größe nach absteigend gezählten m Eigenwerten von K und mit u_i als den zugehörigen orthonormierten Eigenvektoren gilt

$$KU = U \operatorname{diag}(\lambda_i)$$

woraus

$$K = U \operatorname{diag}(\lambda_i) U' \quad \text{und} \quad K^{-1} = U \operatorname{diag}(\lambda_i)^{-1} U'$$

folgt. U bezeichnet die Matrix, deren Spalten aus den Eigenvektoren u_i gebildet wird, und $\operatorname{diag}(\lambda_i)$ die Diagonalmatrix mit den Eigenwerten λ_i in der Diagonale. Wegen der Orthonormalität der Eigenvektoren ist

$$U' = U^{-1}.$$

Die durch Transformation gewonnenen Meßwerte

$$(4.2.3-4) \quad y_1 := \text{diag}(\lambda_i^{-1/2}) U' (x_1 - x)$$

sind dann unkorreliert und vergleichbar (4/15).

Mit (4.2.3-4) berechnet sich die gewöhnliche euklidische Distanz gemäß

$$\begin{aligned} \|y_1 - y_j\|_2^2 &= (y_1 - y_j)' (y_1 - y_j) \\ (4.2.3-5) \quad &= (x_1 - x_j)' U \text{diag}(\lambda_i)^{-1/2} \text{diag}(\lambda_i)^{-1/2} U' (x_1 - x_j) \\ &= (x_1 - x_j)' K^{-1} (x_1 - x_j). \end{aligned}$$

Hieraus läßt sich unschwer eine Distanzfunktion der Art

$$\begin{aligned} d_{K^{-1}}(x_1, x_j) &:= \|x_1 - x_j\|_{K^{-1}} \\ (4.2.3-6) \quad &:= |(x_1 - x_j)' K^{-1} (x_1 - x_j)|^{1/2} \end{aligned}$$

bilden. (4.2.3-6) wird MAHALANOBIS-Distanz genannt. Sie ist skaleninvariant und eliminiert etwaige Korrelationen zwischen den Variablen.

(4.2.3-6) ist ein Spezialfall der Klasse der metrischen, quadratischen Distanzfunktionen

$$(4.2.3-7) \quad d_B(x_1, x_j) := |(x_1 - x_j)' B (x_1 - x_j)|^{1/2}$$

mit B als positiv-definiter Matrix.

Eine weitere Verallgemeinerung stellt die bereits erwähnte L_r -Norm oder L_r -Distanz dar

$$\begin{aligned} d_r(x_i, x_j) &:= \|x_i - x_j\|_r \\ (4.2.3-8) \quad &:= \left| \sum_{l=1}^m |x_{il} - x_{jl}|^r \right|^{1/r} \end{aligned}$$

Von praktischer Relevanz ist neben $r=2$ (euklidische Distanz) der Fall $r=1$.

$$d_1(x_i, x_j) = \sum_{l=1}^n |x_{il} - x_{jl}|$$

Diese Distanz wird auch 'City-block'-Metrik oder 'Taxifahrer'-Metrik genannt (4/16).

Der Grenzfall $r \rightarrow \infty$ hingegen ist von geringerer praktischer Bedeutung, da - wie bereits beschrieben - Extremwerte zu stark gewichtet werden.

4.2.4 DISTANZFUNKTIONEN BEI KLASSEN

Eine Distanzfunktion $d: E \times E \rightarrow R_+$ lässt sich zu einer Distanzfunktion $d: E^* \times E^* \rightarrow R_+$ erweitern, durch die je zwei Teilmengen $g_p, g_q \subseteq E$ eine nicht-negative reelle Zahl $d(g_p, g_q) = d_{pq}$ zugeordnet wird.

Ist die Distanzmatrix $D = (d_{ij})$ vorgegeben, lassen sich etwa folgende Distanzfunktionen definieren

$$(4.2.4-1) \quad d_{pq} = \min_{i \in g_p, j \in g_q} \{d_{ij}\}$$

$$(4.2.4-2) \quad d_{pq} = \frac{1}{n_p n_q} \sum_{i \in g_p} \sum_{j \in g_q} d_{ij}$$

$$(4.2.4-3) \quad d_{pq} = \sum_{i \in g_p} \sum_{j \in g_q} d_{ij}$$

$$(4.2.4-4) \quad d_{pq} = \max_{i \in g_p, j \in g_q} \{d_{ij}\}.$$

Ebenso bestimmt man die Ähnlichkeit zwischen Elementgruppen mit Hilfe von

$$(4.2.4-5) \quad s_{pq} = \min_{i \in g_p, j \in g_q} \{s_{ij}\}$$

$$(4.2.4-6) \quad s_{pq} = \frac{1}{n_p n_q} \sum_{i \in g_p} \sum_{j \in g_q} s_{ij}$$

$$(4.2.4-7) \quad s_{pq} = \frac{1}{n_p n_q} \sum_i \sum_j s_{ij}$$

$$(4.2.4-8) \quad s_{pq} = \max_{i \neq j} \{s_{ij} \mid i \in g_p, j \in g_q\}.$$

Beim Vorliegen quantitativer Variablen wird die Teilmenge g_q E durch ihren Centroid repräsentiert

$$x_{g_p} := \frac{1}{n_p} \sum_i x_i$$

und mittels der euklidischen Distanz bildet man

$$(4.2.4-9) \quad \begin{aligned} d_{pq}^2 &:= \|x_{g_p} - x_{g_q}\|^2 = \frac{1}{n_p n_q} \sum_i \sum_j \|x_i - x_j\|^2 - \\ &\quad - \frac{1}{n_p} \sum_i \|x_i - x_{g_p}\|^2 - \\ &\quad - \frac{1}{n_q} \sum_j \|x_j - x_{g_q}\|^2. \end{aligned}$$

Die Erweiterung der MAHALANOBIS-Distanz zu einer Distanz kann durch

$$(4.2.4-10) \quad d_{pq}^2 := (x_{g_p} - x_{g_q})' K^{-1} (x_{g_p} - x_{g_q})$$

$$(4.2.4-11) \quad d_{pq}^2 := \frac{1}{n} (x_{g_p} - x_{g_q})' W^{-1} (x_{g_p} - x_{g_q})$$

$$\text{mit } W := \sum_{l=1}^k \sum_i \sum_{g_l} (x_i - x_{g_l})(x_i - x_{g_l})'$$

$$(4.2.4-12) \quad d_{pq}^2 := \frac{1}{n_p + n_q} (x_{g_p} - x_{g_q})' W_{pq}^{-1} (x_{g_p} - x_{g_q})$$

$$\text{mit } W_{pq} := W_p + W_q \text{ und } W_p := \sum_i \sum_{g_p} (x_i - x_{g_p})(x_i - x_{g_p})',$$

(analog: W_q)

erfolgen. Je nach Untersuchungsfragestellung wird man eine der vorgeschlagenen Definitionen bevorzugen (4/17).

4.3 VERFAHREN

4.3.1 SYSTEMATISIERUNGSKRITERIEN

Eine Systematisierung von Verfahren zur Numerischen Klassifikation läßt sich nach verschiedenen Kriterien vornehmen (4/18). Generell kann man hierarchische von nicht-hierarchischen Verfahren trennen. Während durch nicht-hierarchische Verfahren eine Repräsentation der Teilmengen von E auf einem in der Regel vorher festgesetzten Zerlegungsniveau erstellt wird, führen hierarchische Verfahren zu einer auf verschiedenen Distanz- oder Ähnlichkeitsebenen anzuordnenden Zerlegung von E und damit zu einem Ergebnis, das anschaulich in Form eines Dendrogramms darstellbar ist. Diese indizierte Hierarchie erhält man entweder durch sukzessive Zerlegung von E oder durch schrittweise Fusion einzelner Elemente zu immer umfassenderen Teilmengen (agglomerative versus divisive Verfahren).

Des weiteren lassen sich Klassifikationsverfahren hinsichtlich der Art der konstruierten Teilmengen unterscheiden, die disjunkt (elementfremd) oder nicht-disjunkt sein können. Meist wird jedes Element einer bestimmten Teilmenge zugewiesen (exhaustive Verfahren), wohingegen die Existenz nicht-klassifizierbarer Elemente vergleichsweise selten vorkommt (nicht-exhaustive Verfahren) (4/19).

Wie auch sonst üblich, kann man Klassifikationsverfahren in iterative und nicht-iterative Verfahren einteilen. Außerdem werden sequentielle von globalen Verfahren unterschieden. Sequentielle oder auch partielle Verfahren gelangen durch Berechnung der oft paarweise definierten Distanz zur gesuchten Klasseneinteilung, während globale Verfahren die Distanz aller Elemente berücksichtigen. Es bietet sich in diesem Fall an, eine geeignete Zielfunktion zu definieren und diese zu minimieren.

Im weiteren Verlauf werden übersichtsartig die im Kontext der Biosignalanalyse wichtigsten Verfahren dargestellt. Es handelt sich ausschließlich um polythetische, d.h. multivariate Verfahren, da monothetische Verfahren, die pro Zerlegungsschritt nur eine Variable betrachten, praktisch und theoretisch von untergeordneter Bedeutung sind.

4.3.2 HIERARCHISCHE VERFAHREN

4.3.2.1 AGGLOMERATIVE VERFAHREN

Agglomerative Verfahren gehen im Unterschied zu divisiven Verfahren von der feinsten Partition aus. Die für zahlreiche agglomerative Verfahren grundlegenden Schritte lassen sich allgemein angeben (4/20):

- (1) Beginne mit der feinsten Partition $G := \{g_1, \dots, g_n\}$, wobei $g_i = \{e_i\}$.
- (2) Suche diejenigen Gruppen, die unter allen übrigen Gruppen die minimale Distanz besitzen, somit die Gruppen g_p und g_q mit $d_{pq} = \min_{i \neq j} d_{ij}$.
- (3) Fusioniere die Gruppen g_p und g_q zur neuen Gruppe g_q^{neu} , wodurch sich die Gruppenanzahl um eins erniedrigt $g_q^{\text{neu}} := g_p \cup g_q$.
- (4) Ändere die q-te Zeile und Spalte der Distanzmatrix, indem die Abstände zwischen der neuen Gruppe g_q^{neu} und allen übrigen Gruppen neu berechnet werden, und streiche die p-te Zeile und Spalte.
- (5) Beende nach n-1 Schritten, wenn also alle Gruppen in einer einzigen Gruppe zusammengefaßt sind, ansonsten fahre bei (2) mit der geänderten Distanzmatrix fort.

Der Unterschied zwischen den verschiedenen agglomerativen Verfahren bezieht sich auf die Definition der Distanz der neu gebildeten Klasse zu den restlichen Klassen (Schritt 4). Damit

lassen sich die relevanten agglomerativen Verfahren über die Rekursionsformel (4/21)

$$(4.3.2.1-1) \quad d_{qi}^{\text{neu}} = \alpha_p d_{pi} + \alpha_q d_{qi} + \beta d_{pq} + \gamma |d_{pi} - d_{qi}|$$

durch Variation der Parameter α_p , α_q , β und γ darstellen.

In der nachfolgenden Tabelle sind die wichtigsten der sich daraus ergebenden Verfahren zusammengestellt.

Verfahren	Parameter				Distanz d_{qi}^{neu}	Eigenschaften/Bemerkungen ($n := n_p + n_q$)
	α_p	α_q	β	γ		
1 Single Linkage (Nearest Neighbor)	1/2	1/2	0	-1/2	$\min(d_{pi}, d_{qi})$	kontrahierend, monoton, invariant unter monotonen Transformationen der Distanzmatrix, Kettenbildung
2 Complete Linkage (Furthest Neighbor)	1/2	1/2	0	1/2	$\max(d_{pi}, d_{qi})$	dilatierend, monoton, invariant unter monotonen Transformationen der Distanzmatrix, neigt zu kleinen Gruppen
3 Average Linkage	1/2	1/2	0	0	$\frac{1}{2}(d_{pi} + d_{qi})$	konservativ, monoton
4 Average Linkage (weighted)	$\frac{n_p}{n}$	$\frac{n_q}{n}$	0	0	$\frac{1}{n}(n_p d_{pi} + n_q d_{qi})$	konservativ, monoton, ge- wichtet mit der Elementzahl
5 Median	1/2	1/2	-1/4	0	$\frac{1}{2}(d_{pi} + d_{qi}) - \frac{1}{4}d_{pq}$	konservativ, nicht monoton (Inversionen), nur für quadierte euklidische Distanzen
6 Centroid	$\frac{n_p}{n}$	$\frac{n_q}{n}$	$-\frac{n_p n_q}{n^2}$	0	$\frac{1}{n}(n_p d_{pi} + n_q d_{qi}) - \frac{n_p n_q}{n^2} d_{pq}$	konservativ, nicht monoton (Inversionen), nur für quadierte euklidische Distanzen
7 Ward	$\frac{n_p + n_i}{n + n_i}$	$\frac{n_q + n_i}{n + n_i}$	$-\frac{n_i}{n + n_i}$	0	$\frac{1}{n + n_i}((n_i + n_p) d_{pi} +$ $+ (n_i + n_q) d_{qi} - n_i d_{pq})$	konservativ, monoton, Minimierung des Varianz- kriteriums bei quadrierten euklidischen Distanzen
8 Flexible Strategie	α	α	$1 - 2\alpha$	0	$\alpha(d_{pi} + d_{qi}) + (1 - 2\alpha) d_{pq}$	$\alpha \rightarrow 0$ kontrahierend, $\alpha = 1/2$ konservativ $\alpha > 1/2$ dilatierend Gunstig: $0.6 \leq \alpha \leq 0.7$.

Abb. 4.3.2.1-1 ÜBERSICHT ÜBER DIE WICHTIGSTEN AGGLOMERATIVEN VERFAHREN

4.3.2.2 DIVISIVE VERFAHREN

Im Unterschied zu den agglomerativen Verfahren gehen divisive Verfahren von der größten Partition aus und verfolgen gleichsam den umgekehrten Konstruktionsweg.

Da divisive Verfahren wesentlich mehr Rechenzeit beanspruchen, wird an dieser Stelle lediglich ein divisives Verfahren vorgestellt, dessen agglomeratives Pendant im Ward-Algorithmus (Verfahren 7) zu sehen ist und dessen Analogie zu dem nachfolgend beschriebenen Verfahren zur Minimierung einer Zielfunktion deutlich ist.

Als Kriterium der "Heterogenität" zwischen den Teilmengen wird das bekannte Varianzkriterium gewählt, was der Forderung

$$\frac{n_1 n_2}{n_1 + n_2} \|x_{g_1} - x_{g_2}\|^2 \rightarrow \max \quad (i \neq j)$$

entspricht.

Anstelle der totalen Enumeration der $2^{n_1 + n_2 - 1} - 1$ Partitionen zur Ermittlung des Extremums wird man im konkreten Fall das einfache Verfahren bevorzugen, welches obige Forderung approximativ erfüllt (4/22):

- (1) Man bestimme in E diejenigen Elemente x_1 und x_2 , die voneinander den größten euklidischen Abstand haben, und bilde $g_1 = \{x_1\}$, $g_2 = \{x_2\}$. Man berechne die Schwerpunkte x_{g_1} und x_{g_2} .
- (2) Suche dasjenige Element aus der Restmenge $E - g_1 - g_2$, welches zu einem der beiden Schwerpunkte der Gruppen g_1 oder g_2 den kleinsten Abstand hat. Man füge es der entsprechenden Gruppe g_i ($i = 1$ oder 2) zu und modifiziere den Schwerpunkt.
- (3) Fahre bei (2) fort oder beende, wenn $E = g_1 \cup g_2$.

Obwohl hierarchische Verfahren (insbesondere der Ward-Algorithmus) sich in der psychologischen Forschung aufgrund vorhandener Software großer Beliebtheit erfreuen, sollte kurz auf grundsätzliche Nachteile hingewiesen werden, die darin bestehen, daß bei ihrer Konstruktion eine (starre) hierarchische Folge von Gruppierungen durchlaufen wird. So ist es keineswegs plausibel, jeweils nur 2 Klassen zu einer neuen zu vereinigen bzw. eine Klasse in zwei aufzuteilen, ohne jedoch die auf dieser Stufe existierende Partition insgesamt zu betrachten. Diese Schwierigkeiten werden (ohne erhöhten Rechenaufwand) durch die im folgenden skizzierten Verfahren überwunden.

4.3.3 VERFAHREN ZUR VERBESSERUNG EINER ZIELFUNKTION

4.3.3.1 EINFÜHRUNG

Das Grundprinzip der im folgenden skizzierten Verfahren besteht darin, daß sie von einer vorgegebenen Anfangspartition G_0 ausgehend eine Folge von Partitionen $G_1, G_2, \dots$ konstruieren, deren Elemente hinsichtlich einer vorher geeignet definierten Zielfunktion $z(G)$ miteinander verglichen werden. Es ist nahelegend, die reellwertige Funktion $z(G)$ so zu definieren, daß sie für eine (aus inhaltlichen oder statistischen Annahmen) besonders gute Partition ein Extremum annimmt. Da nur endlich viele Partitionen G von E existieren, gibt es ein globales Extremum, welches die beste Partition repräsentiert. Unter Punkt (4.1.2) wurde bereits verdeutlicht, daß die Methode der totalen Enumeration schon für kleine Fallzahlen nicht realisierbar ist. Die hier skizzierten heuristischen Verfahren zur Verbesserung einer Zielfunktion ermöglichen daher keine Entscheidung darüber, ob das globale Extremum erreicht ist, vielmehr sind sie in dem Sinne lokal optimal, daß durch den Austausch einzelner Elemente keine weitere Verbesserung erzielt werden kann.

Aus diesem Grunde dürfte es vorteilhaft sein, im konkreten Anwendungsfall diese heuristischen Verfahren mit verschiedenen Startnäherungen durchzurechnen ^(4/23).

Im folgenden wird eine Übersicht über mögliche Zielfunktionen und in der Praxis wichtige Verfahren gegeben. Der Schwerpunkt wird auf die Darstellung des Grundprinzips der Verfahren gelegt, während eine Diskussion der jeweiligen Eigenschaften hinsichtlich Konvergenz, Optimalität, Implementierung oder Anwendbarkeit ausgespart bleibt (4/24).

4.3.3.2 ZIELFUNKTIONEN

Die Funktion $z(G)$ wird man je nach Anwendungsziel verschieden bestimmen. Die hier angegebenen Zielfunktionen sind Beispiele für praktisch relevante Definitionsmöglichkeiten.

Die Zielfunktion

$$z_1(G) := \sum_{l=1}^k \sum_i \left\| x_i - x_{g_l} \right\|^2, \quad (4.3.3.2-1)$$

$$x_{g_1} := \frac{1}{n} \sum_i x_i$$

mit dem Kriterium

$$z_1(G) \rightarrow \min$$

wird Varianzkriterium genannt.

Es gilt, wie aus der Varianzanalyse bekannt, das Theorem:

$$T = W(G) + B(G)$$

mit den "STREUUNGSMATRIZEN"

$$T := \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})'$$

$$W := \sum_{l=1}^k \sum_i (x_i - x_{g_l})(x_i - x_{g_l})'$$

$$B := \sum_{l=1}^k n_l (x_{g_l} - \bar{x})(x_{g_l} - \bar{x})',$$

wobei W und B abhängig von G sind, und es folgt für (4.3.3.2-1)

$$z_1(G) = \text{Spur}W(G) .$$

Daher bezeichnet man (4.3.3.2-1) auch als 'Spur- W '-Kriterium.

Da mit T auch $\text{Spur}T$ unabhängig von G ist und aus dem genannten Theorem

$$\text{Spur}(T) = \text{Spur}(W(G)) + \text{Spur}(B(G))$$

folgt, ist das Varianzkriterium äquivalent zu dem Kriterium

$$\text{Spur}B(G) \rightarrow \text{Max} .$$

Ordnet man das Element x aus der Klasse g_i der Klasse g_j zu und bildet so aus der Partition G die Partition $\tilde{G}$, so ist

$$z_1(\tilde{G}) < z_1(G)$$

genau dann, wenn

$$\frac{n_j}{n_j+1} \|x - x_{g_j}\|^2 < \frac{n_i}{n_i-1} \|x - x_{g_i}\|^2$$

wie man durch elementare Rechnung verifiziert.

Die Testgröße $\Lambda = \det(W)/\det(T)$ dient zur Prüfung der Hypothese, ob die zu den k Teilmengen g_1 ($1 \leq l \leq k$) gehörigen Erwartungsvektoren x_{g_l} gleich sind (4/25). Diese Überlegung führt zur Zielfunktion

$$(4.3.3.2-2) \quad z_2(G) = \det(W(G))$$

mit dem Kriterium

$$z_2(G) \rightarrow \text{Min} ,$$

welches auch Determinantenkriterium genannt wird.

Äquivalent ist die Zielfunktion

$$\begin{aligned}\hat{z}_2(G) &= \det(T)/\det(W) \\ &= \prod_{i=1}^m (1 + \lambda_i(G))\end{aligned}$$

mit dem Kriterium: $\hat{z}_2(G) \rightarrow \text{Max}$.

$\lambda_1(G) \geq \lambda_2(G) \dots \lambda_m(G)$ bezeichnen hierbei die Eigenwerte von $W^{-1}(G)B(G)$ (4/26). $\hat{z}_2(G)$ wird auch Wilks-Lambda-Kriterium genannt.

Ebenso wie (4.3.3.2-2) läßt sich die Zielfunktion

$$(4.3.3.2-3) \quad z_3(G) = \text{Spur}(W^{-1}(G)B(G))$$

mit dem Kriterium

$$z_3(G) \rightarrow \text{Max} ,$$

welches "Spur($W^{-1}B$)"-Kriterium oder einfach "Spur"-Kriterium heißt (4/27), durch die Eigenwerte $\lambda_i(G)$ von $W^{-1}(G)B(G)$ ausdrücken. Es ist

$$\text{Spur}(W^{-1}B) = \sum_{i=1}^m \lambda_i(G) .$$

Werden die Daten gemäß (4.2.3-4) transformiert, so ergibt sich mit (4.2.3-6) die Zielfunktion

$$(4.3.3.2-4) \quad z_4(G) := \sum_{l=1}^k \sum_i (x_i - \bar{x}_{g_l})' K^{-1} (x_i - \bar{x}_{g_l})$$

mit dem Kriterium

$$z_4(G) \rightarrow \text{Min} ,$$

welches dem Varianzkriterium entspricht, wenn anstelle der euklidischen Distanz die MAHALANOBIS-Distanz mit K benutzt wird.

Für die bereits angegebene L_r -Distanz stellt

$$(4.3.3.2-5) \quad z_5(G) := \sum_{l=1}^k \sum_i \|x_i - c_{g_l}\|_r^r$$

mit

$$z_5(G) \rightarrow \text{Min}$$

eine brauchbare Zielfunktion dar. Der Ausdruck c_{g_l} bedeutet im übrigen ein sinnvoll zu wählendes Klassenzentrum (4/28).

4.3.3.3 SIFT-AND-SHIFT VERFAHREN

Die in diesem Punkt skizzierten iterativen Verfahren werden anschaulich als "Sift-and-Shift" Verfahren bezeichnet, da sie pro Iteration jedes Element auf seine Klassenzugehörigkeit sichten (sift) und - falls $z(G)$ dadurch verbessert wird - in eine andere Klasse verschieben (shift).

4.3.3.3.1 ITERATIVES MINIMALDISTANZVERFAHREN

Die Vorgehensweise dieses relativ einfachen Verfahrens läßt sich in folgende Schritte unterteilen (4/29):

- (1) Anfangspartition vorgeben.
- (2) Berechne Gruppenschwerpunkte (Centroide).
- (3) Verschiebe jedes Element in die Gruppe mit dem ihm im Sinne der euklidischen Distanz am nächsten liegenden Schwerpunkt.
- (4) Fahre bei (2) fort oder beende, wenn kein Element in diesem Durchgang die Gruppe gewechselt hat.

Die durch dieses iterative Minimaldistanzverfahren gebildete Folge von Partitionen

$$G^0, G^1, \dots, G^r, G^{r+1}, \dots \quad \text{mit } G^r := \{g_1^r, \dots, g_k^r\}$$

wird hinsichtlich des Varianzkriteriums nicht schlechter, wie einfach zu zeigen ist (4/30).

4.3.3.3.2 AUSTAUSCHVERFAHREN

In Abgrenzung zu dem Minimaldistanzverfahren erfolgt bei dem Austauschverfahren, welches anschaulich auch 'hill-climbing' Verfahren genannt wird (4/31), der Elementenaustausch im Sinne einer direkten Verbesserung des Varianzkriteriums:

- (1) Anfangspartition vorgeben.
- (2) Berechne Gruppenschwerpunkte.
- (3) Prüfe für jedes Element, ob die Gruppierung im Sinne des Varianzkriteriums dadurch verbessert wird, daß es in eine andere Gruppe verschoben wird. Wenn ja, so verschiebe es in die Gruppe mit der größten Verbesserung und berechne die Gruppenschwerpunkte gemäß (4.3.2.1.2) und (4.3.2.1.3) neu.
- (4) Fahre bei (3) fort oder beende, wenn n mal hintereinander kein Gruppenwechsel erfolgt ist.

Im dritten Schritt wird also für jedes Element $x \in g_i$ geprüft, ob

$$n_j / (n_j + 1) \|x - x_{g_j}\|^2 < n_i / (n_i - 1) \|x - x_{g_i}\|^2$$

für eine andere Klasse g_j ($i \neq j$) gilt. Dadurch wird die Zielfunktion $z_1(G)$ durch jeden Austausch eines Elements echt verbessert. Da es nur endlich viele Partitionen gibt, muß dieses Verfahren nach endlich vielen Schritten (in der Regel genügen 5-7 Durchgänge) zu einem lokalen Minimum führen.

4.3.3.3.3 MINIMALDISTANZVERFAHREN UND AUSTAUSCHVERFAHREN FÜR ANDERE ZIELFUNKTIONEN

Die bisher grob beschriebenen Verfahren bezogen sich auf das Kriterium (4.3.3.2-1). Für andere Zielfunktionen lassen sich

diese Verfahren auch allgemein angeben und entsprechend den resultierenden Eigenschaften miteinander vergleichen.

Für das allgemeine Minimaldistanzverfahren ergibt sich:

- (1) Anfangspartition vorgeben.
- (2) Berechne Gruppenzentrum, stelle Distanzfunktion auf.
- (3) Verschiebe jedes Element in die Gruppe mit dem im Sinne der Distanzfunktion am nächsten liegenden Gruppenzentrum.
- (4) Fahre bei (2) fort oder beende, wenn in diesem Durchgang kein Element die Gruppe gewechselt hat.

Demgegenüber kann das Austauschverfahren allgemein beschrieben werden durch:

- (1) Anfangspartition vorgeben.
- (2) Berechne Gruppenzentren.
- (3) Prüfe für jedes Element, ob sich die Zielfunktion dadurch verbessern läßt, daß es in eine andere Gruppe verschoben wird. Wenn ja, so verschiebe es in die Gruppe mit der größten Verbesserung und berechne für die so entstehende Gruppierung die Gruppenzentren neu.
- (4) Fahre bei (3) fort oder beende, wenn n mal hintereinander kein Gruppenwechsel erfolgt ist.

Die Konvergenz- und Optimalitätseigenschaften beider Verfahren sind je nach implementierter Zielfunktion zum Teil recht unterschiedlich und dürften unschwer aus den bisher skizzierten Begriffen und Definitionen abzuleiten sein ^(4/32). Ebenso dürfte man das Austauschverfahren für beliebige Distanzen mit Hilfe der bereit gestellten Informationen entwickeln und programmieren können ^(4/33).

4.3.4 ANDERE VERFAHREN

4.3.4.1 Q-ANALYSE

Analog dem Prinzip der Faktorenanalyse zur Strukturierung von Variablen lassen sich Elemente in bestimmte Klassen oder Gruppen einordnen. In diesem Fall spricht man in Abgrenzung zur R-Technik von der Q-Faktorenanalyse oder kurz Q-Analyse ^(4/34), deren Grundprinzip wie folgt beschrieben werden kann.

Durch Subtraktion der Elementmittelwerte und Normierung, wobei die euklidische Norm zugrundegelegt wird, erhält man aus der Datenmatrix $X = (x_{ij})$ die Matrix

$$Z := (z_{ij}) \quad (i=1,2,\dots,n; j=1,2,\dots,m)$$

mit

$$z_{ij} = x_{ij} / \left(\sum_{j=1}^m \hat{x}_{ij}^2 \right)^{1/2}$$

und

$$\hat{x}_{ij} = x_{ij} - \frac{1}{m} \sum_{j=1}^m x_{ij}$$

und damit die n-reihige Element-zu-Element Korrelationsmatrix

$$Q = ZZ'.$$

Die m-reihige, symmetrische und positiv-semidefinite Matrix

$$T = Z'Z,$$

(die nicht zu verwechseln ist mit der Variablen-zu-Variablen Korrelationsmatrix von X) hat wegen der Normierung von Z den Rang $k \leq m-1$ und entsprechend viele Eigenwerte $\lambda_1 > 0$ zu ortho-normalen Eigenvektoren y_1 ($l=1,2,\dots,k$):

$$Ty_1 = \lambda_1 y_1$$

$$y_1' y_j = \delta_{ij}$$

Dann sind die Vektoren p_1 mit

$$p_1 := \lambda_1^{-1/2} z y_1$$

orthonormale Eigenvektoren zu Q zu den Eigenwerten λ_1 , denn es gilt

$$\begin{aligned} Q p_1 &= Z Z' \lambda_1^{-1/2} z y_1 \\ &= \lambda_1^{-1/2} Z T y_1 \\ &= \lambda_1^{-1/2} Z \lambda_1 y_1 \\ &= \lambda_1 (\lambda_1^{-1/2} z y_1) \\ &= \lambda_1 p_1 \end{aligned}$$

und

$$\begin{aligned} p_1' p_j &= \lambda_1^{-1/2} \lambda_j^{-1/2} y_1' z' z y_j \\ &= \lambda_1^{-1/2} \lambda_j^{-1/2} y_1' T y_j \\ &= \lambda_1^{-1/2} \lambda_j^{-1/2} \lambda_j y_1' y_j \\ &= \lambda_1^{-1/2} \lambda_j^{-1/2} \lambda_j \delta_{1j} = \delta_{1j} . \end{aligned}$$

Faßt man die Vektoren y_1 bzw. p_1 zu den Matrizen Y bzw. P zusammen, so erhält man obige Beziehungen in Matrixschreibweise

$$P = Z Y \operatorname{diag}(\lambda_1^{-1/2})$$

$$P' P = I, Y' Y = I$$

und daraus

$$P \operatorname{diag}(\lambda_1^{1/2}) = Z Y .$$

Nun sind $A := P \text{diag}(\lambda_1^{1/2})$ gerade die Hauptkomponentenladungen von Q , die sich durch obige Identität leicht aus den Eigenvektoren Y von T und aus der transformierten Datenmatrix Z berechnen lassen.

Hieraus ergibt sich das folgende einfache Verfahren der Q -Analyse.

- (1) Berechne Z sowie $T = Z'Z$.
- (2) Berechne Eigenwerte λ_1 und Eigenvektoren y_1 von T und stelle die Matrix Y spaltenweise aus den nach der Größe absteigenden Eigenvektoren zusammen.
- (3) Berechne die $k - m - 1$ gewünschten Hauptkomponenten $A := ZY$ (4.4.1.1)
- (4) Rotiere die Hauptkomponenten A zur Einfachstruktur und erhalte Faktorenladungen.

4.3.4.2 KONFIGURATIONSFREQUENZANALYSE

Im Unterschied zu den bisherigen Verfahren zur Klassifikation quantitativer Daten wird in diesem Punkt ein Ansatz dargestellt, bei dem qualitative Daten zugrundeliegen und dessen Nähe mit den speziell in der Medizin häufig benutzten Strategien zur Analyse mehrdimensionaler Kontingenztabellen offensichtlich ist (4/35).

Bei diesem von LIENERT (4/36) entwickelten und Konfigurationsfrequenzanalyse (KFA) genannten Verfahren bilden mehrkategoriale Daten den Ausgangspunkt.

Bezeichnet r_j die Anzahl der Kategorien der j -ten Variablen, so existieren

$$k = \prod_{j=1}^m r_j$$

Merkmalskombinationen oder Zellen (4/37).

Werden die n Elemente auf die bei m Merkmalen möglichen k Gruppen verteilt, die hier 'Konfigurationen' genannt werden, ergibt sich pro Konfiguration eine bestimmte Frequenz f . Aus der Randsummenverteilung lassen sich für diese beobachteten Häufigkeiten die zugehörigen erwarteten Frequenzen e ermitteln, die zur Berechnung der χ^2 -Prüfgröße benötigt werden. Unter der Nullhypothese der gegenseitigen Unabhängigkeit der Variablen wird eine signifikant überfrequentierte Merkmalkombination als 'Konfigurationstyp' und eine signifikant unterfrequentierte Zelle als 'Konfigurationsantityp' bezeichnet. Wird die totale Unabhängigkeit der Variablen vorausgesetzt, läßt sich die Wahrscheinlichkeit p_1 für die Konfiguration g_1 als das Produkt der Einzelwahrscheinlichkeiten, mit denen ein Element der einzelnen Kategorie $1_1, \dots, 1_m$ ($1 \leq 1_j \leq r_j$) angehört, berechnen:

$$p_1 = \prod_{j=1}^m p_{1,j} .$$

Dabei ist $p_{1,j}$ die Wahrscheinlichkeit des Auftretens der Kategorie 1_j in der j -ten Variablen. Hierbei sind die Einzelwahrscheinlichkeiten aus den Stichproben gewonnene Schätzwerte.

Die erwarteten Häufigkeiten e_1 ergeben sich hieraus mit

$$e_1 = np_1 .$$

Die Nullhypothese $H_0: e_1 = f_1$ wird dann mit Hilfe der χ^2 -Verteilung getestet:

$$\chi^2 = (f_1 - e_1)^2 / e_1$$

mit 1 Freiheitsgrad.

Ist für die Konfiguration g_1 der betreffende χ^2 -Wert laut festgesetztem Alpha-Risiko signifikant, so liegt, wie man sagt, bei g_1 ein Konfigurationstyp vor, falls $e_1 < f_1$ ist, während $e_1 > f_1$ einen Konfigurationsantityp ergibt. Bei multiplen Vergleichen erhöht sich das tatsächliche Fehlerrisiko gegenüber dem nominalen, so daß eine Alphajustierung nach dem Vorschlag

$\alpha_{\text{just}} = \alpha / \text{Anzahl der untersuchten Konfigurationen}$

vorzunehmen ist.

Die Summe aller χ^2 -Komponenten

$$\chi_G^2 = \sum_{l=1}^k (f_l - e_l)^2 / e_l$$

mit $k - (\sum_{j=1}^m 1_j - m + 1)$ Freiheitsgraden gilt als Test für die klassifikatorische Wirksamkeit der KFA. Die KFA stellt somit ein Beispiel für die nicht-exhaustiven Verfahren dar (4/38).

4.3.4.3 ANALYSE DER PUNKTDICHTE

Im Vergleich zu den zahlreichen weiteren vorgeschlagenen Klassifikationsverfahren (4/39) dürfte das auf einem Ansatz von SCHNELL (4/40) basierende Verfahren noch eine gewisse praktische Bedeutung erlangen.

Faßt man die Vektoren x_i ($1 \leq i \leq m$) als Erwartungswerte einer m -dimensionalen Normalverteilung $N(x_i, s)$ mit der Dichtefunktion

$$f(x, x_i, s) = (s^2 2\pi)^{-\frac{m}{2}} \exp(-(x - x_i)'(x - x_i) / (2s^2))$$

auf, überlagern sich diese n Funktionen und die Maxima der Gesamtfunktion

$$F(x, s) = \sum_{i=1}^n f(x, x_i, s)$$

liegen im Punkthaufen.

Das Problem der Klasseneinteilung wird hier auf das Aufsuchen der Maxima der Funktion F zurückgeführt.

Verständlicherweise lassen sich anstelle der Normalverteilung auch andere Belegungsfunktionen, etwa die Funktion

$$h(x, x_i, s) = (1 + s^2 \|x - x_i\|^2)^{-1}$$

oder sogenannte Deltasplines verwenden (4/41).

Darüberhinaus existieren Strategien, die auf der Analyse von Mischverteilungen beruhen (4/42). Gemeinsam ist diesen Ansätzen, daß sie bisher noch keine realisierbare Alternative zu den oben skizzierten Verfahren darstellen.

Aufgrund der sich abzeichnenden Entwicklungstendenzen auf dem Gebiet der numerischen Klassifikation (4/43), welche nicht zuletzt in Abhängigkeit zu dem erreichbaren technischen Fortschritt stehen, sind jedoch in näherer Zeit neue Ansätze, Verfahren und Algorithmen zu erwarten, die dem nicht nur im Kontext der psychophysiologischen Signalverarbeitung gestiegenen Bedarf an effektiven Klassifikationsstrategien entgegenkommen.

Für die Analyse und Klassifikation psychophysiologischer Signale nehmen Fragen der Bestimmung einer angemessenen Zielfunktion und der Auswahl eines geeigneten Verfahrens eine besondere Bedeutung ein, ebenso wie Fragen der Beurteilung der erstellten Klassifikation von weitreichender Bedeutung für die inhaltliche Interpretation der Ergebnisse sind (4/44).

Diese durch eine Anwendung von Klassifikationsverfahren auftauchenden Fragen werden in dem anschließenden Punkten in Beziehung gesetzt zu den beschriebenen Ansätzen der Signalbeschreibung im Zeit- und Frequenzbereich mit dem Ziel, die vorgegebene Thematik dieser Arbeit zu konkretisieren und die sich stellenden Probleme paradigmatisch und in einer möglichst anschaulichen Weise in Angriff zu nehmen.

5 EXEMPLIFIKATION - ANALYSE UND KLASSIFIKATION EINES BEISPIELDATENSATZES AUS DER BIOFEEDBACK-FORSCHUNG

5.1 EINFÜHRUNG

Die in den vorangegangenen Punkten eingeführten Begriffe und Definitionen zur Signalanalyse und zur numerischen Klassifikation werden im folgenden gleichsam paradigmatisch auf einen speziellen psychophysiologischen Datensatz bezogen und hinsichtlich mehrerer Kriterien, insbesondere ihrer Leistungsfähigkeit und Aussagerelevanz genauer diskutiert.

Nach einer eher fragmentarischen Explikation und Erörterung von Aspekten des hier zugrundeliegenden Forschungsprojekts, welches die Problematik der "Stabilisation der Herzrate" zum Gegenstand hat (Punkt 5.2), werden diesem Themenkomplex zuzuordnende und in ihrer formalen Struktur als kennzeichnend ausgewiesene Untersuchungsfragestellungen aufgegriffen (Punkt 5.3) und anschließend mittels ausgesuchter mathematisch-statistischer Verfahren und Modelle bearbeitet. Nach Beschreibung wesentlicher Schritte der numerischen Auswertung (Punkt 5.4) werden die erzielten Resultate auf dem Hintergrund methodischer und substanzwissenschaftlicher Gesichtspunkte konkreter erörtert. Im Mittelpunkt der Betrachtungen finden sich nicht einzelne psychophysiologische Phänomene oder isolierte Prozesse, sondern die aus der beispielhaft durchgeführten Datenauswertung für eine erfolgsversprechende, integrative Analyse psychophysiologischer Daten abzuleitenden Folgerungen (Punkt 5.5).

5.2 ASPEKTE DER BIOFEEDBACK-FORSCHUNG

Die in dieser Arbeit analysierten Daten entstammen jenem Forschungsbereich, der die "wohl faszinierendste Anwendung des Biofeedbacks" (5/1) zu seinem Gegenstand hat. Nach einer ersten und groben Übersicht über diesen derart beurteilten Bereich der "Selbstkontrolle des Herzrhythmus" mit Hilfe von Methoden der biologischen Rückmeldung fällt zunächst auf, daß Studien über die Richtungsänderung der Herzrate (Senkung oder Steigerung) relativ häufig

im Vergleich zu Arbeiten über die Stabilisation der Herzrate (Verringerung der Variabilität) sind ^(5/2).

Von dort nimmt bereits das hier zugrundeliegende Forschungsprojekt eine besondere Stellung ein. Im Zentrum des Interesses steht das Problem, inwieweit die Selbstkontrolle über den aktuellen Verlauf der Variabilität der Herzrate durch differente Formen des Biofeedbacks trainiert und verbessert werden kann. Von hervorragender Bedeutung ist daher die Frage, in welchem Ausmaß subjektiv favorisierte Kontrollstrategien zur Stabilisation der Herzrate identifizierbar und damit therapeutisch nutzbar sind und welche physiologischen, aber auch psychologischen Variablen sich in Relation zu den postulierten Regulationsprozessen methodisch nachweisbar verändern ^(5/3). Da Aussagen der Probanden über realisierte Kontrollstrategien wenig zuverlässig sind, wird besonderes Gewicht auf die Analyse der simultan sich ändernden physiologischen, speziell cardio vasculären Regulationsparameters gelegt und eine ausreichend funktional nutzbare Konfiguration der involvierten Größen im Sinne eines relativ konsistenten "Musters" gesucht. Hinreichend exakte und reliable Befunde stehen auf diesem Gebiete allerdings noch aus ^(5/4).

Zur Beantwortung dieses hier nur kurz anzudeutenden Fragenkomplexes wurde in dem angegebenen Forschungsprojekt den Versuchspersonen ihre eigene Herzrate auf dem Bildschirm eines Oszilloskops zurückgemeldet und in einer Serie von Versuchen wurden die Versuchsbedingungen, welche die Form der Rückmeldung betrafen, systematisch variiert.

Den insgesamt 135 Versuchspersonen, die keine cardio vasculären Beschwerden hatten ^(5/5), wurde die Aufgabe gestellt, ihre eigene Herzrate entsprechend den gegebenen Feedback-Bedingungen möglichst gut an eine zugleich auf dem Bildschirm sichtbare Kriteriumslinie (Referenzlinie) anzunähern. Je geringer die beobachteten Abweichungen waren, desto besser gelang es also den Probanden, ihre Herzrate zu stabilisieren.

EKG, EMG und Atmung wurden mit Hilfe eines Physiopolygraphen (Siemens Mingograph) registriert ^(5/6). Zur Speicherung und Verarbeitung wurden die Meßwerte an einen Prozeßrechner (IBM 1130) weitergeleitet. Die Zeitreihe der aktuellen Herzrate wurde ermittelt, indem nach Transformation der R-Spitzen des nach EINTHOVEN II abgeleiteten EKGs in Rechteckimpulse durch den Zeitzähler eines Interface (WDV) die RR-Intervalldistanzen in Hundertstel-Sekunden berechnet wurden.

Die mittels eines Thermistors aufgenommene Atemtätigkeit wurde von dem Physiopolygraphen in Form von Analog-Spannungsdifferenzen registriert und über den AD-Wandler des Interface dem Rechner übermittelt. Synchron mit der Abtastung der RR-Intervalle wurden die digitalisierten Atemwerte abgespeichert, um eine eindeutige Zuordnung von Herzaktivität und Atemzyklizität zu erreichen. Die Atemamplituden wurden zum Zwecke der Vergleichbarkeit standardisiert.

Die so gewonnene Kurve der Herzrate wurde nach geeigneter Transformation auf einen vor der Versuchsperson positionierten Bildschirm ausgegeben, auf dem neben einer seitlich angeordneten Frequenzskala eine je nach Versuch verschieden definierte Referenzlinie eingeblendet wurde ^(5/7).

Der Versuchsablauf bestand für jeden Probanden aus einer anfänglichen Ruhepause von fünf Minuten sowie sechs fünfminütigen Trainingsphasen, die jeweils durch eine Pause von einer Minute Dauer unterbrochen wurden.

In einer Übersichtstabelle sind die verschiedenen Arten der visuellen Rückmeldung zusammengefaßt, die in den betreffenden Einzelstudien dieses Forschungsprojekts untersucht worden sind ^(5/8).

Nicht zuletzt aus didaktischen Erwägungen sind in dieser Arbeit nur drei spezielle Versuchssituationen von Interesse. Zum einen werden die Daten all jener Personen genauer analysiert, die ein kontinuierliches Feedback erhielten (= Gruppe A). Zum anderen werden die Probanden berücksichtigt, denen alternierend Rückmel-

dung gegeben wurde (Gruppe B =: Feedback in den Trainingsphasen II, IV, VI und kein Feedback während der Phasen I, III, V). Die Mitglieder der Kontrollgruppe erhielten über alle Trainingsphasen keine Rückmeldung, wurden jedoch aufgefordert, ruhig zu sitzen, sich zu entspannen und möglichst regelmäßig zu atmen (=: Gruppe C).

5.3 SKIZZIERUNG DER UNTERSUCHUNGSFRAGESTELLUNGEN

Aus der Vielzahl möglicher Untersuchungsfragestellungen, die zu einem nicht geringen Teil bereits in den angegebenen Publikationen behandelt und intensiver erörtert worden sind ^(5/9), wird jener enger einzugrenzende Bereich ausgewählt, der in Bezug auf die Art der erhobenen Information eine vergleichsweise anschauliche Diskussion der sich stellenden methodischen Probleme erlaubt.

Im Mittelpunkt steht zunächst die Frage, auf welche Weise die komplex vorliegenden Daten über den Verlauf der Herzrate - ohne auf wesentliche Information verzichten zu müssen- reduziert und übersichtlich dargestellt werden können.

Neben Ansätzen und Methoden der Zeitbereichsdarstellung, die in der Regel die Berechnung bekannter statistischer Kenngrößen und geeignet definierter Variabilitätsmaße sowie deren graphische Wiedergabe zumeist in Form von speziellen Histogrammen einschließen ^(5/10), bieten sich hierzu vor allem Methoden der Frequenzbereichsanalyse an.

Ausgangspunkt bilden die Meßwerte, welche die RR-Distanzen in 10^{-2} Sekunden angeben. Durch Bildung des reziproken Werts erhält man in Bezug auf eine feste Zeitbasis die momentane Herzfrequenz. Die Folge dieser Werte, die als "Momentanherzfrequenz" bezeichnet wird ^(5/11), läßt sich aufspalten in einen eher zeitkonstanten Anteil, der im Frequenzbereich von 0.01- 1Hz zu lokalisieren ist und die eigentliche Herzfrequenz angibt, und einen kurzfristig zeitvariablen Anteil, welcher die im psychophysiologischen und medizinisch-klinischen Sinne relevante Herzfrequenz-Variabilität

oder Herzfrequenz-Arrhythmie ausmacht.

Im Unterschied zu anderen psychophysiologischen Signalen ermöglicht nicht das Leistungsspektrum, sondern das Amplitudenspektrum der Momentanherzfrequenz eher aufschlußreiche Interpretationen. Im Amplitudenspektrum der Momentanherzfrequenz sind deutlich ausgeprägte Maxima zu erkennen, die auf unterschiedliche Einflußgrößen und Regulationsmechanismen des kardiovaskulären Systems zurückgeführt werden können (5/12). Im Frequenzbereich < 0.1 Hz dominieren Durchblutungs- und Temperaturrehythmen.

Bei etwa 0.1 - 0.2 Hz treten Blutdruckwellen auf, die zu dem absoluten Maximum des Spektrums führen (5/13).

Von besonderem Interesse in diesem Kontext ist das Frequenzband von 0.1 - 0.4 Hz, in dem die durch die Atemrhythmizität verursachten periodischen Schwankungen sichtbar werden (5/14).

Weitere biochemische, thermische oder hormonelle Parameter haben demgegenüber einen deutlich geringeren Einfluß auf die Gestalt des Amplitudenspektrums und verschwinden zum Teil vollständig im Rauschanteil.

Stellt man diese Eigenschaften des Amplitudenspektrums der Momentanherzfrequenz in Rechnung, speziell, daß sich deutlich erkennbar im Frequenzband 0.1 - 0.4 Hz der Einfluß der Atemrhythmizität ablesen läßt, liegt es in Bezug auf die vorgegebene Qualität der Daten nahe, segmentweise pro Versuchsperson die Amplitudenspektren für die verschiedenen Versuchsphasen zu berechnen und diese dreidimensional darzustellen, um so für den gesamten Versuchsablauf eine bessere Interpretierbarkeit zu erreichen. Da der Atemregulation im Rahmen der Stabilisation der Herzrate eine besondere regulative Funktion zukommt, lassen sich aus dem gezielten intra- und interindividuellen Vergleich der so aufbereiteten Information möglicherweise charakteristische und weitergehende Aufschlüsse und Interpretationsweise über subjektiv differente "Mikro-Strategien" (5/15) zur Stabilisation der Herzfrequenz ableiten.

Hiervon ausgehend ist weiterhin zu fragen, inwieweit aus den in den Amplitudenspektren enthaltenen Informationen eine Klassifika-

tion der Versuchspersonen abzuleiten ist. Man wird von der sicherlich noch expliziter zu begründenden Annahme ausgehen, nach der vor allem Variabilitäten der Atemaktivität einen wesentlichen und unter differenten Bedingungen biologischer Rückmeldung zur Stabilisation der Herzfrequenz nicht zu vernachlässigenden Einfluß auf die Veränderung der Momentanherzfrequenz und damit das Amplitudenspektrum besitzen, wohingegen den im Spektrum enthaltenen weiteren physiologischen Regulationsgrößen eine ausgesprochen geringe Bedeutung hinsichtlich der Experimentalsituation zugeschrieben wird.

In Bezug auf die vorliegenden Meßwerte scheint es daher angezeigt zu sein, speziell die in dem der Atemregulation zuzuordnenden Frequenzbereich auftretenden Kurvenverläufe heranzuziehen und diese numerisch auszuwerten. Aus den verschiedenen Amplitudenspektren wird man also geeignete Parameter extrahieren und die Versuchspersonen auf dieser Basis mittels der im vierten Punkt beschriebenen Ansätze und Verfahren in möglichst intern homogene und extern separierte Cluster einzuteilen versuchen.

Die Frage, welche der berechneten Klassifikationen als "optimal" zu beurteilen ist, wird man nicht eindeutig operational präzisieren und damit beantworten können, wird vorausgesetzt, daß die "wahre", d.h. durch die Untersuchung vorgegebene Information über die Einteilung der PBN in die verschiedenen Klassen (A,B,C) nicht zur Verfügung steht ^(5/16). Mit dem Ziel, die in der Abbildung 1.3-1 wiedergegebenen Schritte der Informationsreduktion unter Bezug auf die bisherigen Ausführungen insbesondere im dritten und vierten Punkt paradigmatisch aufzugreifen und somit die an mehreren Stellen tangierte Problematik des Einsatzes mathematischer und statistischer Verfahren zu konkretisieren, werden die vorliegenden Meßwerte der RR-Distanzen mit Hilfe verschiedener Verfahren aus (4.3.3) und (4.3.4) ausgewertet. Das in (4.1.2) beschriebene Problem, daß clusteranalytische Verfahren trotz ihrer im Vergleich zu anderen multivariaten Ansätzen relativen Nähe zum Merkmalsraum ^(5/17) nur zu sub- oder lokal optimalen Lösungen führen, läßt sich durch einen in dieser Situation nun erst möglichen Vergleich zwischen der experimentell vorgegebenen und

clusteranalytisch definierten Partition veranschaulichen. Neben divergenten Kriterien ^(5/18) liegt es nahe, die empirische und heuristisch erstellte Klasseneinteilung der PBN anhand des folgenden Kriteriums zu vergleichen ^(5/19).

Bezeichne wiederum $G=\{g_1, \dots, g_k\}$ und $G'=\{g'_1, \dots, g'_k\}$ zwei disjunkte und exhaustive Gruppierungen der Menge $E=\{e_1, \dots, e_n\}$, dann ist

$$(5.3.3-1) \quad s(G, G') = \left(\frac{n}{2}\right)^{-1} \sum_{i < j}^n t_{ij} = \frac{2}{n(n-1)} \sum_{i < j}^n t_{ij}$$

mit

$$t_{ij} = \begin{cases} 1 & e_i, e_j \text{ sind in } G \text{ und } G' \text{ in den gleichen Klassen} \\ & \text{oder } e_i, e_j \text{ sind in } G \text{ und } G' \text{ in verschiedenen Klassen} \\ 0 & \text{sonst: } e_i, e_j \text{ sind in } G \text{ in gleichen, aber in } G' \text{ in} \\ & \text{verschiedenen Klassen oder umgekehrt} \end{cases}$$

eine auf der Menge aller Partitionen definierte metrische Ähnlichkeitsfunktion mit $0 \leq s \leq 1$ ^(5/20).

Es wird also gefragt, inwieweit verschiedene Verfahren der numerischen Klassifikation basierend auf den Spektraldaten die latente Strukturierung der PBN aufzudecken in der Lage sind. An den einzelnen Ergebnissen wird man die Abhängigkeit inhaltlich zu interpretierender Resultate von den Prämissen der formalen Konstruktion zumindest partiell ablesen können, und es dürfte dann offensichtlich werden, daß globale Kriterien nur sehr begrenzt zu einer Entscheidung über die zu akzeptierende Lösung beitragen. Aus substanzwissenschaftlichen bzw. psychophysiologischen Erwägungen heraus wird man desweiteren die clusteranalytischen Ergebnisse mit diskriminanzanalytischen Verfahren genauer analysieren wollen. Über eine Aufzeichnung und Interpretation der clusterspezifischen Spektralprofile wird man die Möglichkeit nutzen, in den Spektren jene Parameter oder jenes Frequenzband näher zu identifizieren, das einen statistisch nachweisbaren Diskriminationseinfluß ausübt, was gleichbedeutend damit ist, daß auf diesem Wege relevante Anhaltspunkte über eine Lokalisation der gesuchten interindividuell variierenden und intraindividuell relativ konstanten "Mikro-Kontrollstrategien" zu gewinnen sind ^(5/21).

5.4 SCHRITTE DER NUMERISCHEN AUSWERTUNG

5.4.1 DATENORGANISATION

Angesichts der verfügbaren Rechenkapazität und -zeit ^(5/22) ist es erforderlich, die in den verschiedenen Einzelstudien erhobenen und auf Lochkarten ausgestanzten ca. 2×10^6 Meßwerte ^(5/23) möglichst platzsparend auf Magnetband abzuspeichern. Die Daten der RR-Distanzen, Atemtätigkeit und Phasenkennung werden nach verschiedenen Kriterien auf richtige Reihenfolge, Vollständigkeit und Artefakte kontrolliert. Zur Artefaktselektion werden die Intervallgrenzen vorgegeben, innerhalb derer sich die Meßwerte zu befinden haben. Durch Kummulation der RR-Distanzen werden zudem die durch ihre zeitliche Dauer unterschiedenen Trainings- und Ruhephasen überprüft. Bei Artefakten oder experimentell bedingten Meßpausen wird dann eine Korrektur vorgenommen, wenn die fehlerhaften Werte nicht mehr als 5% des jeweiligen Intervalls betreffen, anderenfalls wird die gesamte Phase bzw. Versuchsperson eliminiert. Durch Berechnung des gleitenden Mittelwerts und der Standardabweichung aus den vorher- und nachgehenden Meßdaten werden mittels eines Zufallsgenerators ^(5/24) die fehlerhaften Werte ersetzt. Stelle und Anzahl der korrigierten Werte werden protokolliert. Überschreitet deren Anzahl bezogen auf das Gesamtintervall die 5% Grenze, wird die Versuchsperson insgesamt nicht weiter berücksichtigt ^(5/25).

Der Datensatz wird völlig neu strukturiert, indem pro PBN eine einheitliche Recordstruktur definiert wird ^(5/26). Diese speziell erzeugte Datei stellt die notwendige Voraussetzung für den im weiteren Verlauf der Berechnungen erforderlichen zeit- und speicherplatzsparenden Zugriff dar. Auf diese Weise wird zudem erreicht, daß in sämtlichen weiteren Programmen dieselbe Subroutine zur Dateneingabe implementiert werden kann ^(5/27). Die zur Datenorganisation geschriebenen PL/1-Programme finden sich unter Punkt

8.1.1 dieser Arbeit.

5.4.2 DATENDESKRIPTION

Im Punkt 3 wurde beschrieben, daß die Berechnung der Fouriertransformation und damit des Spektrums der Amplituden einer Funktion gemäß (3.2.1-6) oder (3.2.1-11) nach verschiedenen Methoden erfolgen kann. Da der FFT-Algorithmus (3.2.3) erst bei Meßwertzahlen >1000 spürbar ins Gewicht fällt (5/28), wird im vorliegenden Fall für $m < 500$ das Amplitudenspektrum durch direkte Berechnung der sin- und cos-Koeffizienten ermittelt (5/29). Da wie auch bei den meisten anderen Verfahren äquidistante Werte vorausgesetzt werden, wird die Zeitreihe der Momentanherzfrequenz entsprechend den geforderten Stützstellen interpoliert. Wie Angaben aus der Literatur und leicht nachvollziehbare eigene Berechnungen ergeben, wird eine hinreichende Genauigkeit schon durch eine lineare Interpolation ersten Grades erzielt (5/30), die zudem von der Rechenzeit her am günstigsten ist.

Mit Hilfe der "Hidden-Line" Technik werden die pro Trainingsphase berechneten Spektren dreidimensional auf einem Plotter (BENSON 10-15) bzw. in einer anderen Programmversion auf einem Bildschirm (TEKTRONIX 4015) ausgegeben, nachdem die Spektraldaten zuvor zwecks Weiterverarbeitung auf Magnetband abgespeichert worden sind.

Die Haupt- und Unterprogramme (FORTRAN IV) zur Berechnung des Amplitudenspektrums und zur dreidimensionalen Darstellung sind im Punkt 8.1.2f. aufgelistet worden.

5.4.3 DATENKLASSIFIKATION

Zur numerischen Klassifikation wird man aufgrund der vergleichsweise einschränkenden Eigenschaften hierarchischer Algorithmen (4.3.2) eines der in (4.3.3) beschriebenen Verfahren zur Verbesserung einer Zielfunktion oder das in (4.3.4.1) erläuterte Verfahren der Q-Analyse heranziehen. Da eine Auswahl und Begründung von mehreren Kriterien abhängig ist, wird die Zahl möglicher Kombinationen von Zielfunktionen und Verfahren nach Plausibilitätsaspekten (5/31) und basierend auf eigenen numerischen Erfahrungen auf folgende Fälle eingeschränkt.

Die im Frequenzbereich 0.01-0.4 Hz erhobenen Meßwerte werden entsprechend dem Ansatz der Q-Analyse geclustert (Strategie A: Q-ANALY) ^(5/32). Derselbe Algorithmus wird auf die zuvor standardisierten Daten ($\bar{x}=0; s=1$) angewandt, um neben einer möglichen Verbesserung - allerdings sehr eingeschränkt - auftretende Effekte einer solchen Merkmalstransformation für diesen speziellen Datensatz grob abschätzen zu können (Strategie B: ZTRANS-QANALY). Desweiteren wird das Austauschverfahren nach dem "hill climbing" Algorithmus (4.3.3.3.2) durchgeführt, das bei gleicher Startpartition in der Praxis häufig zu verbesserten Resultaten als das iterative Minimaldistanzverfahren (4.3.3.3.1) führt (Strategie C: STP-HLCMB). Da eine zufällige Startpartition (STP) ^(5/33) keinerlei Optimalitätseigenschaft bezüglich der implementierten Zielfunktion besitzt, liegt es nahe, zur Verbesserung der Ergebnisse eine geeignete Startpartition zu erzeugen. Im Hinblick auf die vorgegebene Datenqualität dürfte hier die Q-Analyse in Abgrenzung zu den hierarchischen Verfahren eine unter Umständen "bessere" Anfangspartition liefern (Strategie D: QANALY-HLCMB). Geht man von der Annahme ausreichend hoher Meßwertkorrelationen aus, wird man die Daten nach dem Prinzip der Hauptkomponentenanalyse transformieren und im reduzierten Raum eine Klassifikation berechnen. Nach Durchführung einer Hauptkomponentenanalyse (Subroutine KTRANS) ^(5/34) wird eine geeignete Clusterlösung basierend auf dem Austauschalgorithmus im niederdimensionierten Klassifikationsraum gesucht (Strategie E: KTRANS-HLCMB). Für jede der pro Trainingsphase und Strategie gefundenen Klasseneinteilung wird gemäß der Definition in (5.3.3) die Ähnlichkeitsfunktion $s(G, G')$ berechnet.

5.5 DARSTELLUNG UND DISKUSSION DER ERGEBNISSE

5.5.1 DATENDESKRIPTION

Der speziell im dritten Punkt behandelte Übergang von der Zeit- zur Frequenzbereichsdarstellung zielt auf eine adäquate und diagnostisch verwertbare Reduktion der im Zeitbereich noch zu unstrukturierten Information. In den Abbildungen 5.5.1-1/3 sind beispielhaft für die in dem angegebenen Experiment registrierten Meßwerte der Probanden die Amplitudenspektren der Momentanherzfrequenz ab-

Abb. 5.5.1-1

AMPLITUDENSPEKTRUM
 DER MOMENTANHERZFREQUENZ
 (x:= FREQUENZ [Hz];
 y:= AMPLITUDE {normiert};
 z:= PHASE (I-VI))

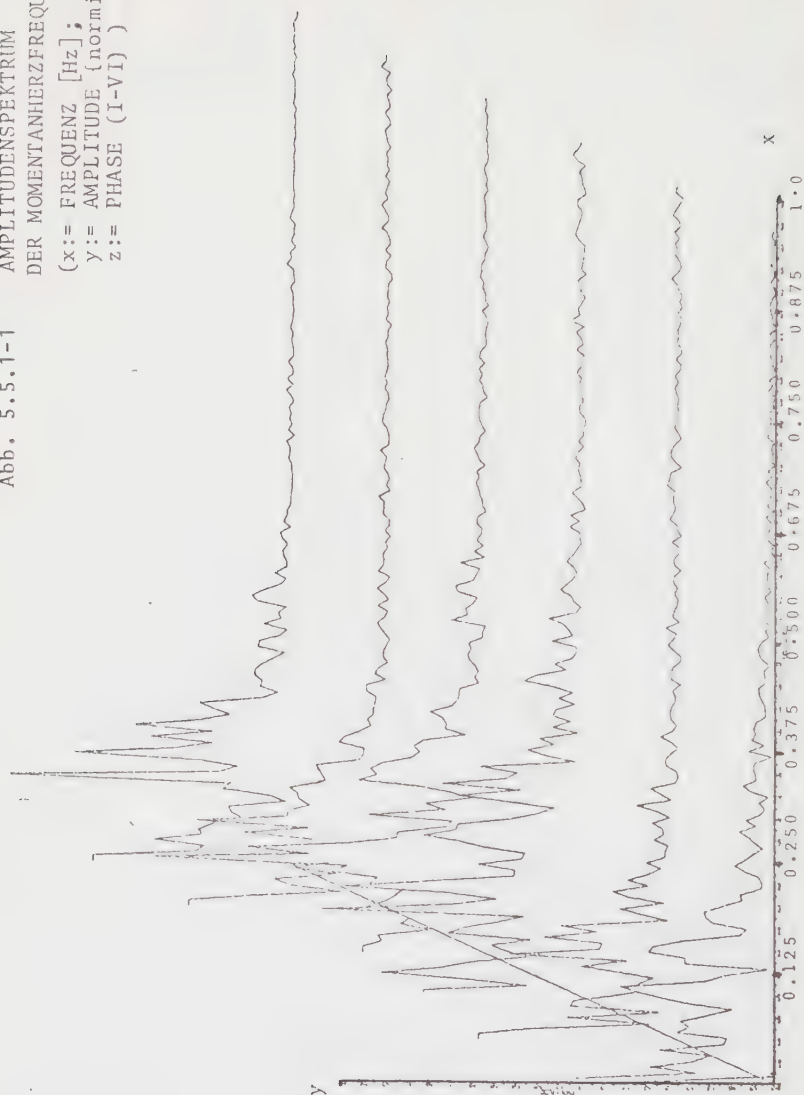


Abb. 5.5.1-2 AMPLITUDENSPEKTRUM
DER MOMENTANHERZFREQUENZ
(x:= FREQUENZ [Hz] ;
y:= AMPLITUDE {normiert} ;
z:= PHASE (I-VI))

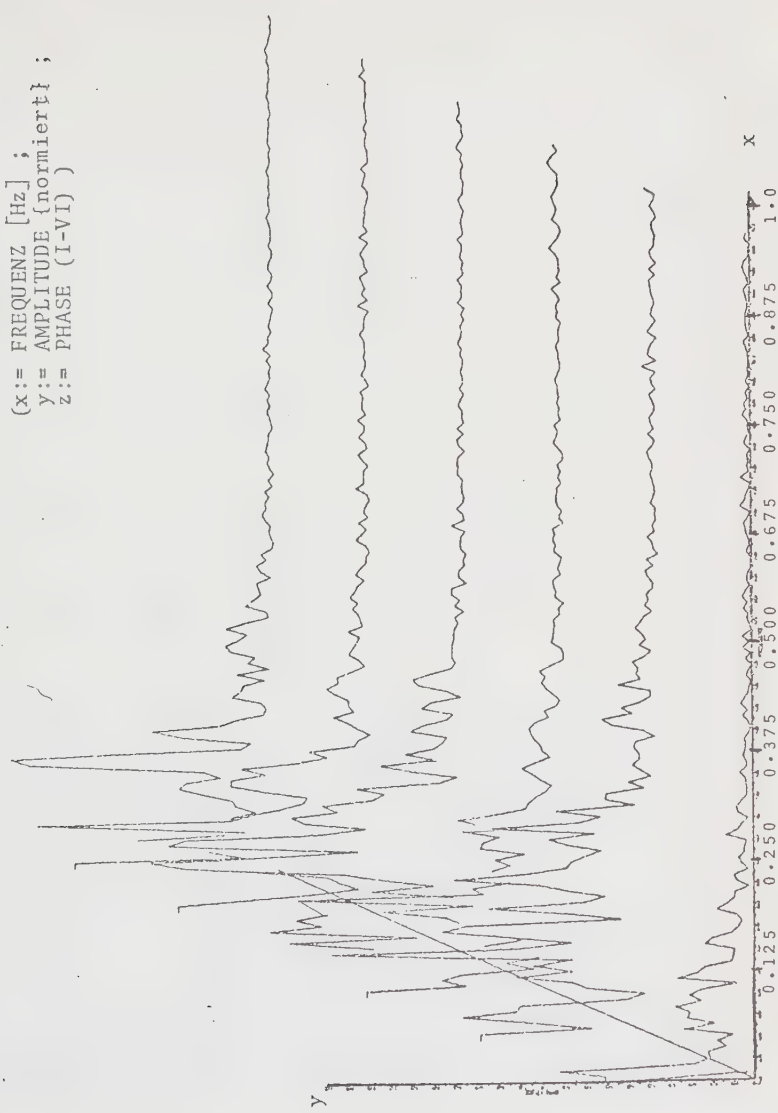


Abb. 5.5.1-3

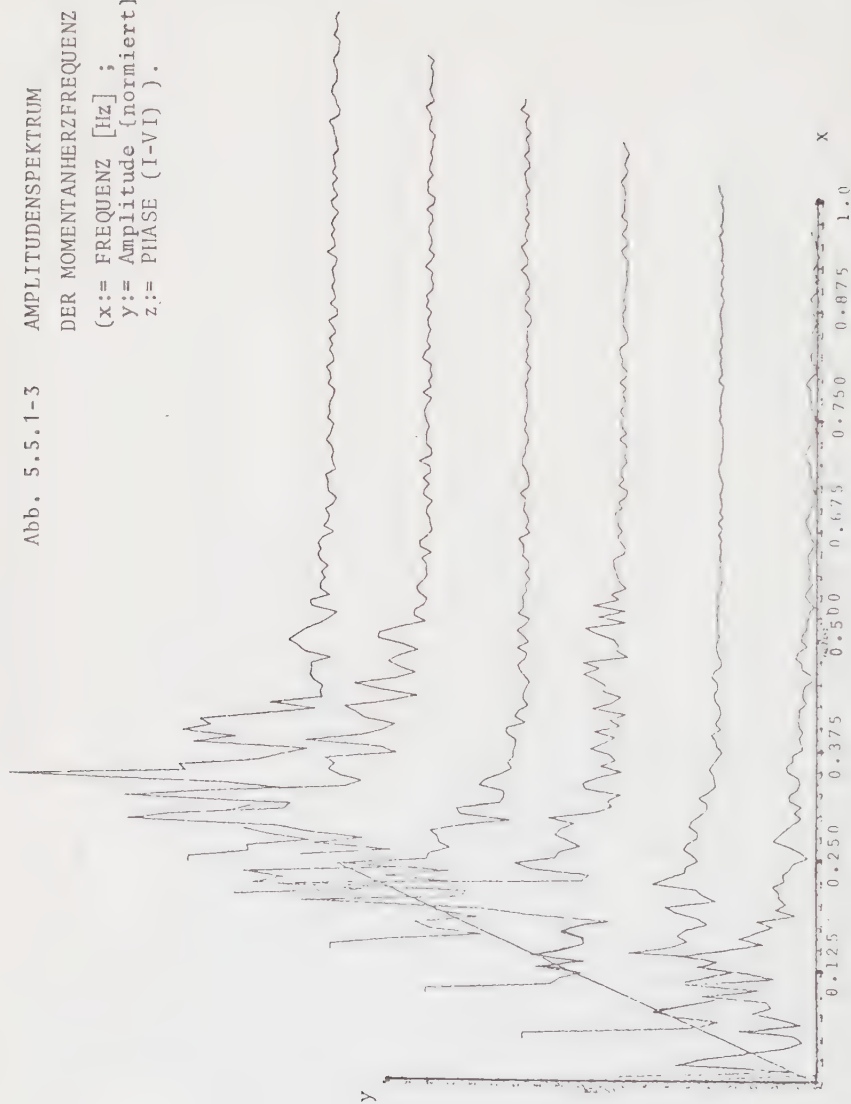
AMPLITUDENSPEKTRUM

DER MOMENTANHERZFREQUENZ

(x:= FREQUENZ [Hz] ;

y:= Amplitude {normiert} ;

z:= PHASE (I-VI)).



gebildet. Die x-Achse markiert in Teilschritten von 0.0125 Hz die Frequenz für das Beobachtungsintervall [0.0125, 1.00 Hz], die senkrecht verlaufende y-Achse gibt den normierten Amplitudbetrag mit $A_i = \{A_i / (A_{\max} - A_{\min})\} \cdot \delta$ an, wobei $A_{\max}$ die maximale, $A_{\min}$ die minimale Amplitude und δ ein geeignet zu wählender, von der Bildgröße bestimmter Skalierungsfaktor ist. Die z-Achse bezeichnet die einzelnen Trainingsphasen (I-VI) des Versuchs.

In den Spektren sind deutlich ausgeprägte und unregelmäßig angeordnete Maxima enthalten, die in Beziehung zu den in (5.3) skizzierten unterschiedlichen cardiovasculären Regulationsgrößen zu setzen sind.

Auffallend ist die Häufung der disproportional großen Amplituden im vorderen Frequenzbereich und das massive Abklingen der Kurve, die im weiteren Verlauf nur noch geringfügig oszillierend sich immer mehr einer einfachen Geraden annähert. Diese Gestalt der Spektren unterstützt die bisher noch nicht belegte, qualitative Aussage, nach der charakteristisch lokalisiert im Amplitudenspektrum der Momentanherzfrequenz deutliche Maxima zu verzeichnen sind, während der übrige Spektralanteil im Sinne eines weißen Rauschens begriffen werden kann.

Ein auf visueller Basis vorgenommener intra- oder interindividueller Vergleich der insgesamt 6x41 Spektren scheitert daran, daß die sehr detaillierten Differenzen nicht nach einheitlichen und operationalisierbaren Kriterien in verschiedene Gruppen oder Klassen eingeteilt werden können, wie bereits durch eine gegenüberstellende Betrachtung der Abbildungen 5.5.1-1/3 ersichtlich sein wird. Die Variabilitäten im Spektrum der Momentanherzfrequenz, die äußerlich zu stark zerklüfteten Kurvenverläufen führen, weisen also keine optisch entschlüsselbaren, intra- oder interindividuell homogenen Strukturen auf, so daß der Untersucher auf eine weiterführende und nach objektivierten, quantitativen Kriterien ausgerichtete Analyse des Datenmaterials angewiesen bleibt.

5.5.2 DATENKLASSIFIKATION

Werden die nach den in (5.4.3) beschriebenen Clusterstrategien berechneten Lösungen zunächst nach formalen Aspekten beurteilt, ergeben sich zwischen den einzelnen Partitionen nicht uninteressante Differenzen, wie der Tabelle 5.5.2-1 zu entnehmen ist, in der die betreffenden Kriteriumswerte gemäß (5.3.3-1) bezogen auf die jeweilige Strategie (A-E) und experimenteller Trainingsphase (I-VI) aufgeführt sind. Die unter der Bezeichnung SQUER abgedruckten Werte stellen den jeweils zeilen- bzw. spaltenweise gebildeten arithmetischen Mittelwert dar.

KRITERIUMSWERTE (5.3.3-1)

	I	II	III	IV	V	VI	Mittelwert
A. CANALY	0.689	0.663	0.705	0.555	0.640	0.620	0.645
B. 2TRANS=>CANALY	0.656	0.707	0.676	0.673	0.623	0.598	0.655
C. STP=>HLCmH	0.645	0.645	0.692	0.590	0.599	0.657	0.639
D. CANALY=>HLCmH	0.759	0.740	0.795	0.759	0.711	0.773	0.763
E. 4TRANS=>HLCmH	0.739	0.753	0.785	0.711	0.766	0.739	0.750
SQUER (I-VI)	0.698	0.711	0.730	0.678	0.686	0.677	

Tab. 5.5.2-1 KRITERIUMSWERTE (5.3.3-1) FÜR DIE CLUSTER-STRATEGIEN (A-E) PRO EXPERIMENTELLER PHASE (I-VI) (SQUER:= SPALTEN- BZW. ZEILENWEISE BERECHNETER ARITHMETISCHER MITTELWERT)

Über alle Strategien (A-E) und Trainingsphasen (I-VI) hinweg ist zu erkennen, daß je nach analysierter Phase die Abweichungen zwischen der experimentell bestimmten, in diesem Fall latenten und der numerisch ermittelten Klasseneinteilung der Elemente derart variieren, daß die Unterschiede zwischen den verschiedenen Clusterstrategien im Mittel etwa doppelt so groß sind wie diejenigen zwischen den einzelnen Trainingsphasen ($s_{\Delta} = 0.118$; $s_{\Delta} = 0.53$).

Der spaltenweise Vergleich ergibt, daß die Klassifikation durchschnittlich für die Phase III im obigen Sinne am besten gelingt, wohingegen die Cluster der Phase VI die geringsten Ähnlichkeiten mit der experimentellen Einteilung der Versuchspersonen zu den angegebenen Gruppen (A,B,C) besitzen.

In Abgrenzung zu diesen relativ gering ausgeprägten Unterschieden ergeben sich bei dem zeilenweisen Vergleich für die verschiedenen Clusterstrategien (A-E) deutlichere Differenzen. Die pro Zeile berechneten Werte belegen die auch methodisch begründbare Annahme, nach der die jeweilige Startpartition einen nicht unwesentlichen Einfluß auf das Klassifikationsergebnis besitzt. Zu betonen ist, daß die jeweiligen Lösungen nur lokal optimal oder suboptimal sind, wie in (4.1.2) anhand der Stirling'schen und Bell'schen Zahlen einführend aufgezeigt worden ist, was die betreffenden Differenzen verständlich werden läßt.

Die Kombination von Q-Analyse und Austauschverfahren führt durchschnittlich zu den besten Ergebnissen (Strategie D), während mit einer zufälligen Anfangspartition mit demselben Algorithmus erkennbar schlechtere Resultate über alle Phasen hinweg erzielt werden (Strategie C). Wie ebenfalls im vierten Punkt einführend beschrieben, lassen sich diese Differenzen auch dadurch erklären bzw. herleiten, daß die zufällige Anfangspartition keinerlei Optimalitätseigenschaften bezüglich der implementierten Zielfunktion besitzt. Die Gegenüberstellung der durch die Clusterstrategien D und C definierten Partitionen verweist m.a.W. auf den nicht unerheblichen Einfluß der jeweiligen Startpartition auf die Güte der erstellten Klassifikation.

Demgegenüber hat eine Standardisierung der Daten in diesem Fall keinen so klar ausgeprägten Einfluß. Das Verfahren der Q-Analyse erweist sich als relativ unempfindlich oder robust gegenüber einer entsprechenden Merkmalstransformation. Während sich also die Gruppierungen für die nicht standardisierten Daten (Strategie A) und standardisierten Daten (Strategie B) bezüglich der Q-Analyse nur unwesentlich unterscheiden ($s_{\Delta} = 0.1$), fällt insbesondere auf, daß die Transformation der Daten gemäß dem Prinzip der Hauptkomponen-

tenmethode in dieser Situation einen gravierenderen Einfluß zu besitzen scheint.

Wird - um den Einfluß der Startpartition auf das Klassifikationsergebnis auszuschalten - in allen Fällen dieselbe Standardanfangspartition zugrundegelegt, ergibt sich, daß die Anwendung des Austauschverfahrens entsprechend dem "hill-climbing" Algorithmus zu deutlich unterschiedlichen Klasseneinteilungen führt, wenn zum einen die Rohdatenmatrix (Strategie C) und zum anderen die gemäß dem Prinzip der Hauptkomponentenmethode reduzierte Datenmatrix (Strategie E) zur Basis der numerischen Klassifikation benutzt wird. Die durch das Vorgehen in E konstruierten Partitionen liegen im Schnitt um das ca. 6 fache weiter entfernt von den Ergebnissen der Strategie C als von denen der im obigen Sinne definierten besten Strategie D ($s_{E-C} = 0.81$; $s_{D-E} = 0.13$). Die auch inhaltlich nicht uninteressante Bedeutung dieser Unterschiede wird noch dadurch unterstrichen, als durchgängig, d.h. für alle Phasen die durch die Clusterstrategie E erzeugten Partitionen höhere Kriteriumswerte als in C aufweisen.

Dieser bereits formal augenfällige Befund, der auf eine hohe und hinsichtlich der konstruierten Cluster relevante Korrelation der Meßwerte aufmerksam macht, legt die genauere Analyse und Überprüfung der Ergebnisse mit Hilfe des diskriminanzanalytischen Modells nahe, wobei man zweckmäßigerweise anstelle der direkten Berechnung die schrittweise multiple diskriminanzanalytische Auswertung bevorzugen wird.

Von den Ergebnissen der pro Phase für die Clusterstrategie E separat berechneten Diskriminanzanalyse sind hier nur die für den weiteren Untersuchungsgang wichtigen Werte wiedergegeben worden, da eine speziellere psychophysiologische Interpretation aus methodischer Sicht an diesem Punkt sekundär ist. Mit nur geringen und in keinem Fall statistisch überzufälligen Modifikationen können die hier für die Phase IV aufgelisteten Werte auf die anderen Phasen übertragen werden. (5/35).

Neben dem herausragenden Resultat, daß der Reklassifikationsfehler im Mittel (!) kleiner als 3% ist, was ein Maß für die

statistisch hochsignifikante Trennung der gefundenen Cluster ist, fällt ohne auf weitere Details einzugehen auf, daß die Verteilung der WILKS' Lambda bzw. F-Werte kennzeichnende Extrema für die Frequenzen 17-22 besitzt. Analoge Verteilungen liegen für die restlichen Phasen vor, in denen bis auf einzelne Ausreißer (<5%) sich die konstruierten Cluster primär hinsichtlich jener Variablen unterscheiden, die in den Frequenzbereich von 0.11-0.20 Hz einzuordnen sind, wie für die Phase IV aus der Tabelle 5.5.2-2 ersichtlich ist.

Wird also von der Annahme ausgegangen, daß bei einer ungefähren mittleren Frequenz von 0.16 Hz (oder rund 10 Atemzügen pro Minute), welche unter Bedingungen der Entspannung auch physiologisch zu erwarten sein dürfte, die Cluster im diskriminanzanalytischen Sinn optimal getrennt sind ($p < 0.01$), wird die psychophysiologische Hypothese durch diese Befunde belegt und unterstützt, daß Prozessen der Atemtätigkeit innerhalb des kardiovaskulären Systems eine (bisher nicht hinreichende erfaßte) regulative Funktion für eine Stabilisation der Herzfrequenz unter differenten Feedback-Bedingungen zugeschrieben werden kann.

Durch diese unter Ausnutzung der spektralen Information erzielten Ergebnisse wird also belegt, daß Personen unter verschiedenen, experimentell systematisch variierten Bedingungen der biologischen Rückmeldung interindividuell ^(5/36) sehr homogene, klassenspezifische Reaktionsmuster besitzen. Bezogen auf die realisierten Experimente beträgt die Rate der richtig klassifizierten Versuchspersonen mehr als 75%, d.h. es ist in mehr als dreiviertel aller Fälle möglich, aus den Amplitudenspektren der Momentanherzfrequenz die Zugehörigkeit einer Versuchsperson zu einer der drei Versuchsgruppen richtig anzugeben. Auf diesem Hintergrund lassen sich differenziertere Rückschlüsse hinsichtlich der applizierten Feedback-Bedingungen ziehen, denen auch auf diesem Wege eine für die Variabilität der Momentanherzfrequenz entscheidende Rolle eingeräumt wird.

<u>Variable</u>	<u>WILKS'</u> <u>Lambda</u>	<u>F</u>
Freq. 1	0.913	1.809
2	0.668	9.415
3	0.937	1.270
4	0.861	3.058
5	0.884	2.490
6	0.990	0.178
7	0.898	2.155
8	0.942	1.164
9	0.879	2.610
10	0.985	0.273
11	0.991	0.171
12	0.891	2.317
13	0.863	3.003
14	0.872	2.781
15	0.747	6.410
16	0.866	2.923
17	0.520	17.479
18	0.658	9.851
19	0.533	16.617
20	0.485	20.154
21	0.253	55.957
22	0.604	12.441
23	0.993	0.122
24	0.889	2.351
25	0.988	0.213
26	0.780	5.344
27	0.840	3.598
28	0.888	2.383
29	0.983	0.315
30	0.857	3.154
31	0.896	2.198
32	0.955	0.892
33	0.585	13.444
34	0.722	7.282
35	0.931	1.391
36	0.934	1.328
37	0.784	5.220
38	0.920	1.651
39	0.959	0.812
40	0.899	2.120

Tab. 5.5.2-2 WILKS' LAMBDA UND F-WERTE
FÜR DIE CLUSTERSTRATEGIE (E) /
PHASE (IV)

Da es nicht das Ziel einer unter der vorgegebenen Thematik verfaßten Arbeit sein kann, die aus diesen Befunden abzuleitenden und im engeren Sinne psychophysiologisch bedeutenden Konsequenzen auszuarbeiten und genauer darzustellen, soll abschließend eine sich in dieser Situation stellende Auswertungsalternative skizziert werden, die nach methodisch-methodologischen Kriterien angezeigt sein dürfte.

Die durch die Clusterstrategie (E) definierten Partitionen und die Ergebnisse der multiplen Diskriminanzanalyse machen auf die Wichtigkeit der Variablen im Frequenzband 0.11-0.20 Hz für die Klasseneinteilung aufmerksam, so daß es naheliegend ist, anstelle des bisher analysierten umfangreicheren Datenmaterials nur jene Meßwerte zur Klassifikation der Elemente heranzuziehen, die in dem betreffenden Frequenzbereich erhoben worden sind.

Da die Clusterstrategien C und D im Vergleich zur Q-Analyse (Strategie A und B) in der Regel zu besseren Gruppierungen führten, wird analog der Strategie C und D das Austauschverfahren gemäß dem "hill-climbing" Algorithmus für die entsprechend selektierten Variablen berechnet. Anstelle der bisher 40 Variablen werden hier lediglich 11 Variablen zugrundegelegt. Die ermittelten Ergebnisse, die wiederum separat für jede der experimentellen Trainingsphasen (I-VI) berechnet worden sind, finden sich in der folgenden Tabelle 5.5.2-3.

KRITERIUMSWERTE (5.3.3-1)

	I	II	III	IV	V	VI	SQUER
F. STP=>HLCMB	0.746	0.740	0.746	0.702	0.746	0.711	0.732
G. <TRANS=>HLCMB	0.795	0.780	0.759	0.759	0.743	0.746	0.763

Tab. 5.5.2-3 KRITERIUMSWERTE (5.3.3-1) FÜR DIE CLUSTER-STRATEGIEN (F+G) PRO EXPERIMENTELLER PHASE (I-VI) (SQUER:= ZEILENWEISER ARITHMETISCHER MITTELWERT)

Der angestrebte Vergleich zwischen den Tabellen 5.5.2-1 und 5.5.2-3 ergibt, daß die auf der Grundlage von nur 11 Variablen ermittelten Klasseneinteilungen der Elemente hinsichtlich des Kriteriums (5.3.3-1) nicht wesentlich von denjenigen abweichen, die im Falle von 40 Variablen erreicht worden sind. Es wird sogar deutlich, daß dieselbe Clusterstrategie (STP => HLCMB) für $m = 11$ im obigen Sinne bessere Partitionen erzeugt. Hinsichtlich der experimentell vorgegebenen Einteilung der Elemente ist es m.a.W. möglich, basierend auf der Information von lediglich 11 Spektralwerten die Klassenzugehörigkeit einer Versuchsperson mit einer Wahrscheinlichkeit $>75\%$ richtig anzugeben.

Psychophysiologisch dürfte von besonderem Interesse sein, daß die erstellten Cluster statistisch hochsignifikant ($p < 0.01$) verschieden sind, wie aus den Ergebnissen der für diese Partitionen berechneten multiplen, schrittweisen Diskriminanzanalyse abzulesen ist.

Entsprechend der bisher auf formal-methodische Punkte eingeschränkten Darstellung werden auf den folgenden Seiten nur die formal wesentlichen Resultate wiedergegeben. Die in dieser Situation mit der Clusterstrategie (G) für die Trainingsphase (II) beispielhaften Ergebnisse finden sich in den Tabellen 5.5.2-4/6. Wie auch in den anderen Fällen können die hier ermittelten Werte mit nur geringen, keinesfalls gegensinnigen Abänderungen auf die anderen Phasen übertragen werden.

Neben der genaueren Interpretation der diskriminanzanalytischen Resultate wäre aus psychophysiologischer Perspektive speziell die nähere Untersuchung der clusterspezifischen Spektralprofile sinnvoll. Einen ersten, anschaulichen Einblick mag die Abbildung 5.5.2-1 geben, in der die Lokalisation der Clusterelemente in der berechneten Diskriminanzebene eingetragen ist.

<u>Variable</u>	<u>WILKS'</u> <u>Lambda</u>	<u>F</u>
Freq. 1	0.784	5.216
2	0.898	2.151
3	0.923	1.567
4	0.994	0.098
5	0.970	0.580
6	0.942	1.167
7	0.901	2.079
8	0.961	0.771
9	0.952	0.941
10	0.798	4.803
11	0.887	2.400
12	0.999	0.018
13	0.937	1.276
14	0.977	0.437
15	0.923	1.572
16	0.911	1.847
17	0.919	1.674
18	0.530	16.834
19	0.287	47.184
20	0.227	64.374
21	0.319	40.506
22	0.877	2.659
23	0.373	31.832
24	0.211	70.681
25	0.363	33.227
26	0.123	135.471
27	0.583	13.564
28	0.535	16.482
29	0.427	25.407
30	0.459	22.315
31	0.574	14.100
32	0.837	3.676
33	0.876	2.676
34	0.760	5.982
35	0.758	6.049
36	0.792	4.961
37	0.741	6.628
38	0.829	3.917
39	0.736	6.789
40	0.853	3.267

Tab. 5.5.2-4 WILKS' LAMBDA UND F-WERTE
FÜR DIE CLUSTERSTRATEGIE (G) /
PHASE (II)

AMPLITUDENSPEKTREN
DISKRIMINANZANALYSE DER CLUSTER

DISCRIMINANT ANALYSIS									
SUMMARY TABLE									
STEP NUMBER	VARIABLE ENTERED REMOVED	F TO ENTER CR REMOVE	NUMBER INCLUDED	WILKS' LAMBEA	SIG.	RAD*5 V	CHANGE IN RAD*5 V	SIG. OF CHANGE	
1	FFQ01A	135.47168	1	0.12300	0.000	27.94038	27.94038	0.0	
2	FFQ02B	21.35278	2	0.05864	0.000	212.68218	48.74180	0.0	
3	FFQ02B	19.15598	3	0.02801	0.000	424.68722	112.00514	0.0	
4	FFQ02A	12.81456	4	0.01640	0.000	672.13745	247.44912	0.0	
5	FFQ01A	6.03046	5	0.01210	0.000	808.60566	136.46851	0.0	
6	FFQ01B	0.97559	6	0.00081	0.000	890.79492	72.11896	0.0	
7	FFQ02B	3.82163	7	0.00067	0.000	976.67456	85.87964	0.0	
8	FFQ027	2.13640	8	0.00596	0.000	1018.60010	41.92554	0.0	
9	FFQ027	1.54547	9	0.00540	0.000	1081.21216	62.61206	0.0	

DISCRIMINANT FUNCTION	EIGENVALUE	RELATIVE PERCENTAGE	CANONICAL CORRELATION	FUNCTIONS DERIVED	WILKS' LAMBEA	CHI-SQUARE	DF	SIGNIFICANCE
1	21.560770	74.34	0.977	0	0.0054	177.500	10	0.0
2	7.08525	25.66	0.538	1	0.1193	72.700	8	0.000

REMAINING COMPUTATIONS WILL BE BASED ON 2 DISCRIMINANT FUNCTION(S)

STANDARDIZED DISCRIMINANT FUNCTION COEFFICIENTS

	FUNC 1	FUNC 2
FFQ01A	0.15201	0.15204
FFQ02B	0.12600	0.11200
FFQ021	0.05174	-0.15611
FFQ022	-0.00457	-0.11700
FFQ024	0.16077	-0.10675
FFQ025	0.00763	-0.11426
FFQ026	0.04467	0.21120
FFQ027	0.02422	0.06761
FFQ028	-0.16744	-0.14626

AMPLITUDENSPEKTREN
DISKRIMINANZANALYSE DER CLUSTER

CENTROIDS OF GROUPS IN REDUCED SPACE

	FLNC 1	FUNC 2
GROUP 1	0.037186	0.17620
GROUP 2	-0.18012	-1.07026
GROUP 3	1.05650	0.07182

AMPLITUDENSPEKTREN
DISKRIMINANZANALYSE DER CLUSTER

DIRECTORY OF PLOT SYMBOLS

SYMBOL	GROUP	LABEL
1	1	
2	2	
3	3	
U		ALL UNGROUPED CASES
*		CENTROIDS

AMPLITUDENSPEKTREN
DISKRIMINANZANALYSE DER CLUSTER

PREDICTION RESULTS

ACTUAL GROUP	PREDICTED GROUP MEMBERSHIP			NO. OF CASES
	GP. 1	GP. 2	GP. 3	
GROUP 1	26.0 100.0%	0.0 0.0%	0.0 0.0%	26.
GROUP 2	0.0 0.0%	5.0 100.0%	0.0 0.0%	5.
GROUP 3	0.0 0.0%	0.0 0.0%	10.0 100.0%	10.

PERCENT OF "GROUPED" CASES CORRECTLY CLASSIFIED: 100.00%

Tab. 5.5.2-6

ERGEBNISTABELLE (2) DER SCHRITTWEISEN
MULTIPLEN DISKRIMINANZANALYSE FÜR DIE
CLUSTERSTRATEGIE (G)/ PHASE (II)

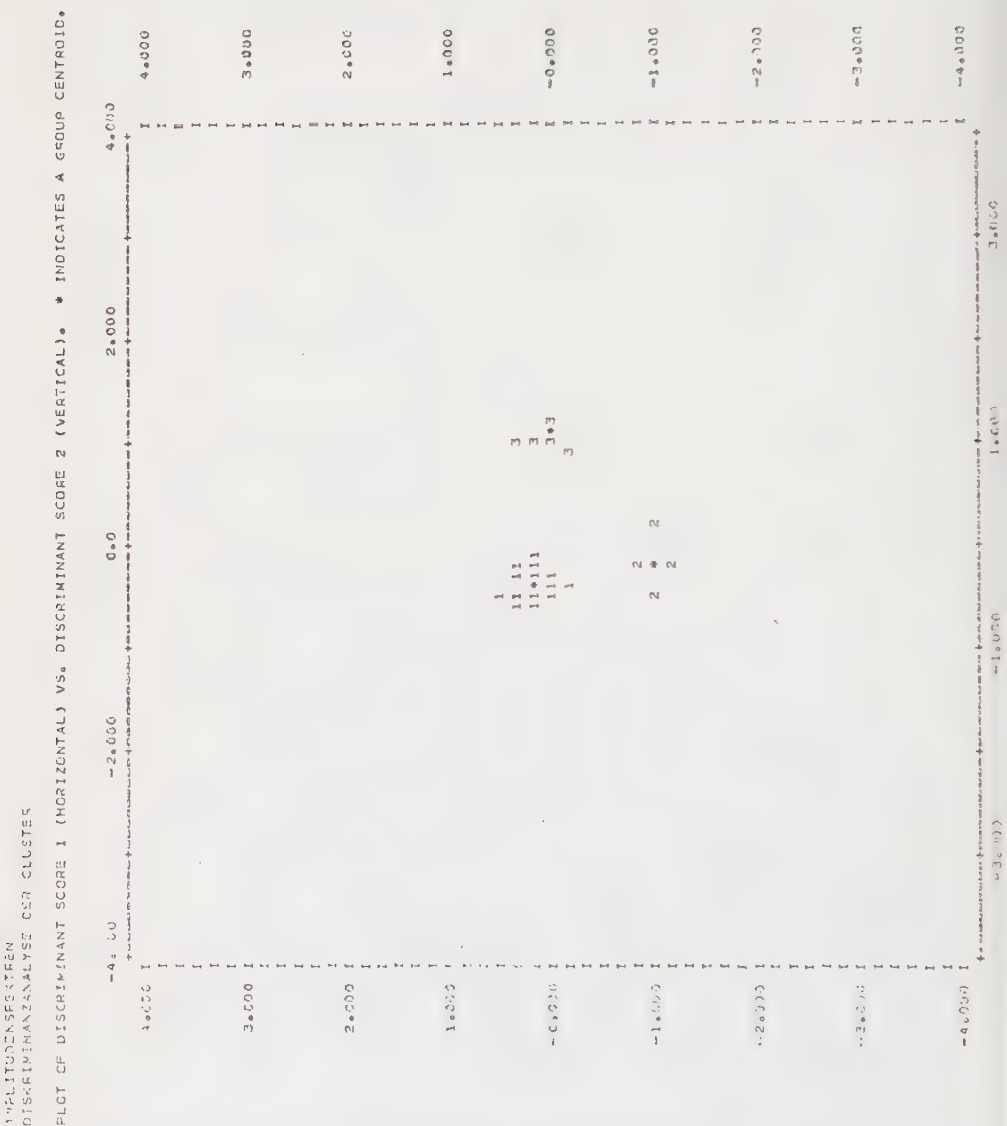


Abb. 5.5.2-1

LOKALISATION DER CLUSTERELEMENTE
IN DER DISKRIMINANZEBOE (CLUSTER-
STRATEGIE (G)/ PHASE (II))

Anstelle einer für die Entwicklung präzisierbarer Hypothesen relevanten, in diesem Kontext aber nachgeordneten genaueren Datenanalyse und -interpretation seien die bisher gewonnenen Ergebnisse gemäß der auf formal-methodische Aspekte eingeschränkten Betrachtungsweise kurz zusammengefaßt.

Aus den Ergebnissen der verschiedenen Clusterstrategien ist keine zwingende Präferenz für einen bestimmten Algorithmus hinsichtlich der Angemessenheit oder Güte, die experimentell bestimmte Klassenstruktur zu identifizieren, abzuleiten, obgleich die Kombination von Q-Analyse und Austauschverfahren im Mittel den übrigen Clusterstrategien überlegen ist. Dieser Effekt wird mit durch die mittels der Q-Analyse erreichte Verbesserung der Startpartition für den "hill-climbing" Algorithmus im Unterschied zur zufälligen Startpartition bedingt sein, welche - wie im vierten Punkt einführend beschrieben - auch keinerlei Optimalitätseigenschaften bezüglich der implementierten Zielfunktion besitzt.

Eine Merkmalstransformation im Sinne einer Standardisierung, durch die unter Umständen die Klassifikation zu verbessern gewesen wäre, hat für den vorliegenden Datensatz keine nachweisbare Bedeutung.

In Abgrenzung zu dieser Robustheit der Q-Analyse gegenüber einer Standardisierung führt die Transformation der Daten nach dem Prinzip der Hauptkomponentenanalyse zu ausgesprochen guten Ergebnissen, sowohl was die Ähnlichkeit gemessen an der Kriteriumsfunktion (5.3.3-1) als auch die diskriminanzanalytische Überprüfung anbelangt.

Dieser Sachverhalt wird auf die Struktur des Datenmaterials, d.h. die ausreichend hohen Meßwertkorrelationen zurückgeführt, was auch dadurch nahegelegt wird, daß eine Separierung der Cluster in der Diskriminanzebene mit einer durchschnittlichen Wahrscheinlichkeit >97% möglich ist.

Die in einem abgrenzbaren Frequenzbereich registrierbare Häufung trennender Variablen, die an den statistisch zum Teil hochsignifikanten WILKS' Lambda oder F-Werten abzulesen ist, führt zur Klassifikation der Elemente basierend auf der Information über dieses spezielle Frequenzband. Es zeigt sich, daß mit diesen sehr wenigen Daten in etwa gleich gute Klassifikationen gefunden werden können; in jedem der Fälle sind die Cluster hochsignifikant separiert.

Aus psychophysiologischer Perspektive ergibt sich der hier nur sehr fragmentarisch in seiner weiteren Bedeutung abzuschätzende Befund, daß in einem relativ exakt abgrenzbaren und primär der Atemregulation zuzuordnenden Frequenzband des Amplitudenspektrums der Momentanherzfrequenz hochsignifikante Differenzen zwischen den clusteranalytisch definierten Gruppen bestehen.

Abgesehen von der an sich nicht uninteressanten Übereinstimmung zwischen der numerisch konstruierten und experimentell bestimmten Einteilung der Versuchspersonen legen die berechneten Daten eine Interpretation nahe, nach der sich im Spektralbereich statistisch hochsignifikant separiert intern sehr homogene und über alle Phasen hinweg relativ konstante Reaktionsmuster auf differente Bedingungen der biologischen Rückmeldung identifizieren lassen.

Sowohl die äußerst geringe Zahl der zur Klassifikation benötigten Variablen als auch die Art der ermittelten Diskriminanzfunktionen weisen darauf hin, daß in dem definierten Frequenzband intensivere Analysen erfolgsversprechend sein dürften.

6 WEITERFÜHRENDE ANALYSESTRATEGIE

6.1 EINFÜHRUNG

Die bisher erzielten Ergebnisse haben Anhaltspunkte über Lokalisation und Art subjektiv realisierter und relativ gruppenhomogener "Mikro-Kontrollstrategien" unter differenten Biofeedback-Bedingungen geliefert. Detaillierte qualitative Befunde konnten aufgrund der vorgegebenen Daten und damit Informationsstruktur nur sehr bedingt im Unterschied zur quantitativ ausgerichteten Interpretation der numerischen Resultate erhoben werden. Angesichts der komplex vermaschten Regulationsvorgänge zwischen der Herz- und Atemtätigkeit, die sich gleichsam mehrfach gebrochen in den ausgewerteten Spektraldaten manifestierten, scheint eine genauere und methodisch gesicherte Analyse der Koppelung von Herz- und Atemfunktionen angezeigt zu sein, die über zum Teil situationsinvariante Phänomene, nach denen mit einer Inspiration immer eine Erhöhung und mit einer Expiration ein Absinken der Herzrate verbunden ist, hinaus eine prozeßorientierte Aufschlüsselung des unter den verschiedenen experimentellen Bedingungen biologischer Rückmeldung ablaufenden psychophysiologischen Geschehens zumindest partiell ermöglicht.

Als eine verständlicherweise in ihrer Aussagekraft und -reichweite relative, dennoch in diesem Problemkontext nützliche Strategie dürfte der im folgenden Text entwickelte analytische Ansatz gelten, in dem versucht wird, ausgehend von einem speziellen Begriffsinstrumentarium aus der mathematischen Theorie der Differentialgleichungen wesentliche der auf der subjektiven Mikroebene anzuordnenden Regulationsvorgänge zu erfassen und nach explizierten Kriterien zu interpretieren. Inwieweit die intendierte Charakterisierung der psychophysiologischen Prozesse im Sinne abgrenzbarer, unter divergierenden experimentellen Voraussetzungen relativ konstanter Funktionszusammenhänge möglich ist, wird nach der Behandlung der mathematischen Aspekte (6.2), welche die Definition eines stabilen Systemzustands einschließt, in einem gesonderten Punkt mit Hilfe von Methoden der graphischen Datenverarbeitung zu beantworten versucht (6.3).

6.2 MATHEMATISCHE ASPEKTE
 6.2.1 GRUNDBEGRIFFE
 6.2.1.1 DIFFERENTIALGLEICHUNGEN

Sind in einer Gleichung neben den unabhängigen Veränderlichen noch die Funktionen dieser Veränderlichen enthalten, bezeichnet man derartige Gleichungen als Funktionalgleichungen. Existieren darüberhinaus noch die Ableitungen dieser Funktionen und gehen sie in die Gleichung ein, spricht man von Differentialgleichungen (6/1).

Gesucht werden in dieser Situation also Funktionen $y=f(x)$ unabhängiger Veränderlicher x . Bei den gewöhnlichen DGLen hängt $f(x)$ nur von einer veränderlichen Größe ab und es gehen nur die Ableitungen nach dieser einen Veränderlichen ein, während bei den partiellen DGLen die gesuchte Funktion $f(x_1, x_2, \dots)$ von mehreren Veränderlichen abhängt und die jeweiligen partiellen Ableitungen in die Gleichung eingehen (6/2).

Eine gewöhnliche DGL läßt sich in allgemeiner Form schreiben als

$$(6.2.1.1-1) \quad F(x, y(x), \dots, y^{(n)}(x)) = 0,$$

wohingegen eine partielle DGL etwa zweier unabhängiger Veränderlicher x_1, x_2 in der Form

$$(6.2.1.1-2) \quad F(x_1, x_2, y(x_1, x_2), y_{x_1}, y_{x_2}, y_{x_1 x_1}, y_{x_2 x_2}, y_{x_1 x_2}) = 0$$

notiert wird. Die Indizes x_1, x_2 bezeichnen die partiellen Ableitungen nach x_1 bzw. x_2 , so z.B.

$$y_{x_1} = \frac{dy}{dx_1} \quad \text{und} \quad y_{x_1 x_2} = \frac{d^2 y}{dx_1 dx_2}.$$

Die Ordnung einer DGL wird durch den höchsten in ihr vorkommenden Differentialquotienten festgelegt.

Die allgemeine gewöhnliche DGL zweiter Ordnung hat also die Gestalt

$$(6.2.1.1-3) \quad F(x, y, y', y'') = 0 \quad .$$

Im Unterschied zu dieser impliziten Schreibweise versteht man unter einer expliziten DGL n-ter Ordnung eine nach der höchsten in ihr enthaltenen Ableitung aufgelöste DGL. Die allgemeine Form einer solchen Gleichung lautet somit

$$(6.2.1.1-4) \quad y^{(n)} = G(x, y, y', \dots, y^{(n-1)}) \quad \text{bzw.} \\ y^{(n)} - G(x, y, y', \dots, y^{(n-1)}) = 0 \quad .$$

Im einfachsten Fall sei z.B. die DGL $y' = y$, dann löst jede Funktion $y = a e^x$ für $a \in \mathbb{R}$ die vorgegebene DGL, d.h. man erhält im allgemeinen eine Lösungsschar. Legt man jedoch zusätzlich einen Funktionswert, den sog. "Anfangswert", fest, so ist die Lösung stets eindeutig bestimmt: für $y(0)=1$ besitzt die obige DGL nur die Lösungsfunktion $y = e^x$.

Die Lösung einer derartigen DGL ist also im allgemeinen eine Schar von Funktionen, so daß weitere Freiheitsgrade existieren und zwar soviele, wie die Ordnung der DGL angibt. Durch Festsetzung von entsprechend vielen Anfangswerten erhält man eindeutige Lösungen unter mathematisch allgemeinen Voraussetzungen. Zur numerischen Lösung dieser Aufgaben kann man verständlicherweise eine Reihe von verschiedenen Methoden heranziehen (6/3).

6.2.1.2 LINEARE DIFFERENTIALGLEICHUNGSSYSTEME UND EIN DIFFERENTIALGLEICHUNGSSYSTEM N-TER ORDNUNG

Als lineares DGL-System bezeichnet man ein System der Form (6/4)

$$(6.2.1.2-1) \quad y'(x) = A(x) y(x) + b(x) \quad ,$$

wobei $y(x) = (\eta_\nu(x))_{(n,1)}$ sowie $b(x) = (\beta_\nu(x))_{(n,1)}$ Spaltenvektoren im $\mathbb{R}^n$ sind und $A(x) = ((\alpha_{\nu\mu}(x)))_{(n,n)}$ eine reell- oder auch komplexwertige (n,n) -Matrix bedeutet.

Die allgemeine lineare DGL n-ter Ordnung hat in Bezug auf die Notierung (6.2.1.1-1) die Form

$$(6.2.1.2-2) \quad \eta^{(n)}(x) + \alpha_{n-1}(x) \eta^{(n-1)}(x) + \dots + \alpha_0(x) \eta(x) = v(x) \quad .$$

$\alpha_0(x), \alpha_1(x), \dots, \alpha_{(n-1)}(x)$ sind hierbei reell- bzw. komplexwertige Funktionen. Der Ausdruck $\eta^{(v)}(x)$ bezeichnet die Ableitung

$$\frac{d^v}{dx^v} \eta \quad .$$

Die spezielle Anfangswertaufgabe ist

$$y'(x) = A_S(x) y(x) + b_S(x)$$

$$y(x_0) = y_0$$

mit

$$A_S(x) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \dots & \dots & \dots & 0 \\ -\alpha_0(x) & \dots & -\alpha_1(x) & \dots & \end{pmatrix} \quad \text{und}$$

$$b_S(x) = (0, \dots, 0, v(x))' \quad \text{mittels}$$

$$y(x) = (\eta_v(x))_{(n,1)} \quad , \quad \text{wobei}$$

$$\eta_1(x) = \eta(x), \eta_2(x) = \eta'(x), \dots, \eta_n(x) = \eta^{(n-1)}(x)$$

äquivalent zu

$$\eta^{(n)}(x) + \alpha_{(n-1)} \eta^{(n-1)} + \dots + \alpha_0 \eta = v(x)$$

$$\eta(x_0) = \eta_{10}, \eta'(x_0) = \eta_{20}, \dots, \eta^{(n-1)}(x_0) = \eta_{n0} \quad ,$$

denn

$$\begin{pmatrix} n' \\ n(2) \\ \vdots \\ n(n) \end{pmatrix} = \begin{pmatrix} 0 & 1 & \dots & 0 \\ 0 & 1 & & \\ \vdots & & \ddots & \\ \alpha_0(x) & \dots & \alpha_{n-1}(x) \end{pmatrix} \begin{pmatrix} n \\ n(2) \\ \vdots \\ n(n-1) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ v(x) \end{pmatrix}.$$

Dieser Zusammenhang ist für die folgende Diskussion von Bedeutung. Man kann sich also auf die Behandlung von DGL-Systemen beschränken und sämtliche Ergebnisse gemäß der oben aufgezeigten Äquivalenz auf die linearen DGLen n -ter Ordnung übertragen.

Die Lösungen eines linearen DGL-Systems stellen einen Raum mindestens der Dimension n dar mit der Basis $v_1, v_2, \dots, v_n$. Man kann so jede Lösung als Linearkombination

$$(6.2.1.2-3) \quad y(x) = \sum_{i=1}^n p_i v_i(x) \quad \text{schreiben.}$$

In Analogie zu den Ausführungen in Punkt 2.5.2.2 bezeichnet man die Basis des Lösungsraums als Fundamentalsystem des DGL-Systems.

Sind in der Matrix A sämtliche Funktionen konstant, also $A = (a_{\nu\mu})$, so spricht man auch von einem linearen DGL-System mit konstanten Koeffizienten, deren Lösungen eine besonders einfache Gestalt haben und auf das Matrizen eigenwertproblem der Matrix A zurückzuführen sind, so daß man auf einfache Weise aus dem Eigensystem der Matrix A ein Fundamentalsystem angeben kann.

Betrachtet man ein homogenes System, d.h. $b(x) = 0$, von gewöhnlichen DGLen erster Ordnung mit konstanten Koeffizienten, läßt sich dieses System in der Matrizenform schreiben als

$$(6.2.1.2-4) \quad y' = A y \quad \text{mit} \quad A = (a_{ij}) \quad (1 \leq i, j \leq n).$$

Setzt man nun

$$y(x) = u e^{\lambda x}, \quad \text{wobei} \quad u \in \mathbb{R}^n, \quad \lambda \in \mathbb{R} \quad \text{ist,}$$

erhält man mit $y'(x) = \lambda u e^{\lambda x}$ nach Division durch $e^{\lambda x}$ für (6.2.1.2-4) die Gleichung

$$(6.2.1.2-5) \quad A u = \lambda u .$$

Diese Gleichung stellt nun die bereits aus der multivariaten Statistik bekannte Eigenwertaufgabe der Matrix A dar. Die λ_i werden wiederum als Eigenwerte von A und die u_i als (zugehörige) Eigenvektoren bezeichnet^(6/5).

Es ist allerdings zu beachten, daß im Gegensatz zur multivariaten Statistik die Matrix A als beliebig und nicht als symmetrisch und positiv-definit angenommen wird, wodurch das Eigensystem eine erheblich kompliziertere Gestalt haben kann und u.U. auch bedeutend schwieriger numerisch zu ermitteln ist.

Ohne auf die daraus resultierenden Komplikationen bezüglich mehrfacher Eigenwerte und zugehörigen Eigen- und Hauptvektoren einzugehen, sei bemerkt, daß das aus dem durch das Eigenwertproblem (6.2.1.2-5) ermittelte Fundamentalsystem gemäß (6.2.1.2-3) zu Lösungen führt, die sich als Linearkombinationen von Funktionen(vektoren)

$$p(x) e^{\lambda x}$$

schreiben lassen, wobei λ ein Eigenwert von A und $p(x)$ ein Vektor ist, dessen Komponenten Polynome in x sind. Für einfache Eigenwerte reduzieren sich im übrigen diese "Polynome" auf Konstanten, d.h. die Komponenten der zugehörigen Eigenvektoren.

6.2.2 STABILE ZUSTÄNDE

6.2.2.1 EINFÜHRUNG

Für die Behandlung der insbesondere in der Psychophysiologie aufgeworfenen Frage der Regelung vorgegebener Systemgrößen ist eine geeignete formale Beschreibung des untersuchten Systems, die u.a. mit Hilfe der bisher skiz-

zierten Begriffe erfolgen kann, unerläßlich. Bei einer solchen Analyse von Regelungsvorgängen, die man nicht selten als DGL-Systeme der involvierten Regelungsgrößen schreiben kann, ist es oft von hervorragendem Interesse, ob sich das betrachtete System in einem Gleichgewicht oder in der Nähe davon befindet.

Aus der Physiologie etwa ist bekannt, daß neuronale-vegetative und humorale Größen so miteinander verknüpft sind, daß das betreffende Teilsystem des Körpers durch geeignete, manchmal minimale Änderung dieser Regelungsgrößen auf ein bestimmtes, zeitlich begrenztes Gleichgewichtsniveau eingestellt werden kann. In dieser Situation hängt also der Zustand des Systems nicht mehr von der Zeit ab. Darüberhinaus ist es wichtig zu wissen, in welchem Ausmaß geringfügige Störungen den erreichten Gleichgewichtszustand beeinflussen.

Bis auf sehr wenige Ausnahmen sind jedoch die Funktionen für die beteiligten Regelungsgrößen, aus denen man sofort den Gleichgewichtszustand und damit die Stabilität des Systems erkennen könnte, nicht bekannt, so daß man auf einen anderen Zugang zur Lösung dieses Problems angewiesen ist. Man geht dabei von den folgenden Definitionen aus^(6/6).

6.2.2.2 DEFINITIONEN

In der Umgebung des festen Punktes x_s sei das System gewöhnlicher DGLen definiert

$$(6.2.2.2-1) \quad x' = \tilde{F}(x).$$

Gilt $\tilde{F}(x_s) = 0$, so heißt x_s singulärer Punkt. Diesen Punkt nennt man im Rahmen der Steuerungs- und Regelungstechnik auch Gleichgewichtslage oder stationären Zustand.

Die konstante und d.h. stationäre Lösung der obigen Gleichung ist also

$$(6.2.2.2-2) \quad x(t) = x_s = \text{konst.}$$

Durch geeignete Transformationen kann man leicht erreichen, daß sich der singuläre Punkt im Nullpunkt des Koordinatensystems be-

findet, was die Diskussion der verschiedenen Fälle stationärer Zustände vereinfacht.

Da die Anwendung der auf LJAPUNOV zurückgehenden Definition der Stabilität eines Systems unter explizierten Voraussetzungen zu aufschlußreichen Ergebnissen in der Praxis führt, wird dieser Ansatz zunächst betrachtet^(6/7).

Die stationäre Lösung x_s eines Systems der Form (6.2.2.2-1) heißt stabil, wenn für jedes $\epsilon > 0$ ein $\delta(\epsilon, t_0) > 0$ derart existiert, daß aus

$$(6.2.2.2-3) \quad \|x(t_0) - x_s\| < \delta(\epsilon, t_0)$$

auch die Bedingung

$$(6.2.2.2-4) \quad \|x(t) - x_s\| < \epsilon \quad \text{für } t > t_0$$

folgt. Ist die Bedingung (6.2.2.2-4) nicht erfüllt, befindet sich also das System in einem instabilen Zustand.

Als Norm $\|\cdot\|$ wird man eine der aus (2.5.2.1) bekannten Funktionen wählen und diese im Sinne einer Distanz interpretieren^(6/8).

In manchen Anwendungssituationen ist die obige Stabilitätsdefinition noch zu schwach und man wird eine strengere Definition zugrundelegen. Man definiert daher den Zustand asymptotischer Stabilität.

Die Lösung x_s eines Systems (6.2.2.2-1) heißt asymptotisch stabil, wenn

(6.2.2.2-5) für x_s Bedingung (6.2.2.2-3/4) gilt und

(6.2.2.2-6) jede in einer gewissen Umgebung von x_s beginnende Bewegung mit $t \rightarrow \infty$ gegen x_s strebt. Also existiere ein festes $\delta(t_0) > 0$ derart, daß $\lim_{t \rightarrow \infty} x(t) = x_s$ für alle Lösungen mit

$$\|x(t_0) - x_s\| < \delta \quad \text{gilt.}$$

Es gilt in diesem Fall dann der Satz (ohne Beweis)^(6/9):

- Wenn alle Eigenwerte der Matrix A des linearen Systems (aus (6.2.1.2-4)) negativen Realteil haben, ist die Null-Lösung

dieses Systems asymptotisch stabil.

Dieser Satz kann auf den Fall nicht-linearer Systeme übertragen werden, insofern diese in einem gewissen Sinne "sehr nahe" bei einem linearen System liegen. Es gilt hier:

- Die Funktion G besitze in der Umgebung des Punktes $x=0$ eine Darstellung der Form $G(x) = Ax + h(x)$, wobei mit $h(x)$ noch $h(x)/||x||$ für $||x|| \rightarrow 0$ gegen 0 strebt. Falls die Realteile aller Eigenwerte von A negativ sind, ist die Null-Lösung von G asymptotisch stabil.

6.2.2.3 FALLUNTERSCHIEDUNGEN UND GRAPHISCHE DARSTELLUNG DER LÖSUNGSKURVEN

Die bisher erörterten Ansätze zur Definition eines stabilen Zustands eines mit Hilfe von DGLen angebbaren Systems werden im folgenden für den einfachen Fall $n=2$ betrachtet. Hier ergeben sich bereits verschiedene Fälle, die eine für praktische Zwecke aufschlußreiche Darstellung der Lösungskurven implizieren.

Im Fall $n=2$, also bei einem linearen DGL-System erster Ordnung liegt das zu untersuchende System in der Form vor

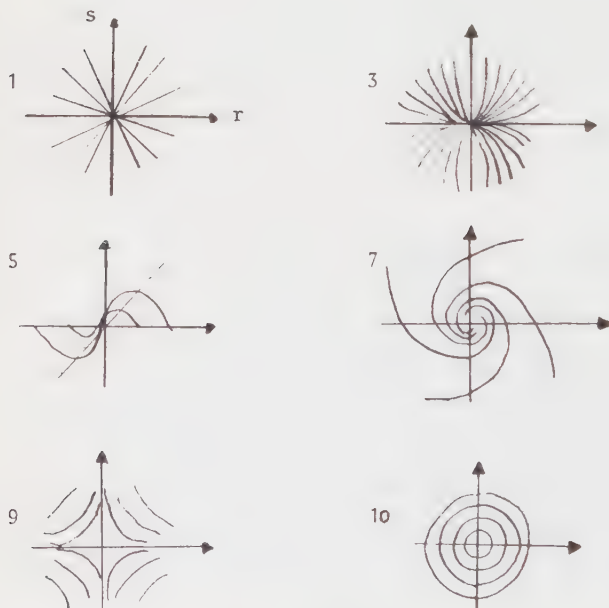
$$(6.2.2.3-1) \quad \begin{aligned} \dot{r} &= a r + b s \\ \dot{s} &= c r + d s \end{aligned} \quad \text{bzw. in Matrixenschreibweise}$$

$$\begin{pmatrix} \dot{r} \\ \dot{s} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} r \\ s \end{pmatrix}$$

Hinsichtlich der beiden Eigenwerte λ_1, λ_2 der Matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ lassen sich unterschiedliche Stabilitätsaussagen über den Punkt $(0,0)$ machen. In der Tabelle 6.2.2.3-1 sind die verschiedenen Fälle sowie die zugehörigen Lösungskurven zusammengestellt (6/10), wohingegen die Lösungskurven graphisch in der Abb. 6.2.2.3-1 wiedergegeben werden.

FALL	EIGENWERTE	LÖSUNGSKURVE
1	negativ reell, $\lambda_1 = \lambda_2$	Knotenpunkt 1.Art, stabiler Stern
2	positiv reell, $\lambda_1 = \lambda_2$	unstabiler Stern
3	negativ reell, $\lambda_1 \neq \lambda_2$	Knotenpunkt 2.Art, stabiler zweitangentiger Knoten
4	positiv reell, $\lambda_1 \neq \lambda_2$	unstabiler zweitangentiger Knoten
5	negativ, $\lambda_1 = \lambda_2$	Knotenpunkt 3.Art, stabiler eintangentiger Knoten
6	positiv, $\lambda_1 = \lambda_2$	unstabiler eintangentiger Knoten
7	$\operatorname{Re}(\lambda) \pm j\operatorname{Im}(\lambda)$ $\operatorname{Re}(\lambda) < 0$; $\operatorname{Im}(\lambda) \neq 0$	stabiler Strudelpunkt
8	wie 7, jedoch $\operatorname{Re}(\lambda) > 0$	unstabiler Strudelpunkt
9	reell, $\lambda_1, \lambda_2 < 0$	Sattelpunkt
10	$j\operatorname{Im}(\lambda)$, $\lambda \neq 0$	Zentrum

Tab. 6.2.2.3-1 FALLUNTERSCHIEDUNG UND LÖSUNGSKURVEN

Abb. 6.2.2.3-1 LÖSUNGSKURVEN IN DER PHASENEBENE (r, s)

Die in der Abbildung aufgezeichneten Kurvenverläufe in der (r,s)-Phasenebene sind zur Tab. 6.2.2.3-1 in Beziehung zu setzen.

Für praktische Anwendungen ist m.a.W. von besonderem Interesse, inwieweit das untersuchte System einen im obigen Sinne definierten stabilen Zustand erreicht hat und sich daher hinsichtlich der Lösungskurven ein stabiler Knoten, Stern, Sattelpunkt oder etwa ein Zentrum ergibt. In welchem Maß eine auf diesem Zugang der Systembetrachtung aufbauende Visualisierung stabiler Zustände und anschließende qualitative Interpretation in der Praxis von Bedeutung sein kann, mag das im nächsten Punkt skizzierte Beispiel verdeutlichen.

6.3 KONKRETISATION

6.3.1 EINFÜHRUNG

In Abgrenzung zu stark formalisierten Arbeiten über ein regelungstheoretisches Kreislaufmodell, mit dem eine zuverlässige Analyse arbeitsphysiologischer und rhythmologischer Einflüsse auf die Momentanherzfrequenz angestrebt wird ^(6/11), ist in dem vorliegenden psychophysiologischen Kontext der Biofeedback-Forschung zur Stabilisation der Herzrate ausschnitthaft die gegenseitige Beeinflussung der Herz- und Atemfrequenzen von Interesse, da von hier aus unter präzisierten Annahmen relevante Schlüsse auf intrasubjektiv anzuordnende Mechanismen kardio-respirodynamischer Kontrollen abgeleitet werden können. Der intendierten Visualisierung des betrachteten Regelungsvorgangs kommt also primär eine hypothesengenerierende Funktion zu.

Da ein hinreichend ausformuliertes DGL-System, das den Zusammenhang der Variablen Herz- und Atemfrequenz wiedergibt, nicht ohne weiteres angegeben werden kann, wird basierend auf den Darstellungen aus (6.2) eine qualitative Analyse der wechselseitigen Änderungsprozesse vorgeschlagen, um von dort aus genügend exakte Orientierungsdaten für die Entwicklung weiterführender Hypothesen zu erhalten.

Es wird hier also angenommen, daß ein valides DGL-System über den

interessierenden psychophysiologischen Prozeß nicht verfügbar ist, die betreffenden Variablen jedoch einem solchen System genügen. Bekannt sind allerdings die Lösungen für die Herz- und Atemfrequenzen, insofern während der Versuche die empirischen Daten aufgezeichnet und abgespeichert worden sind. Aus dieser Information läßt sich in Analogie zu (6.2.2.3) die gesuchte Darstellung im Phasenraum (r,s) konstruieren.

In Bezug auf die in der Tab. bzw. Abb. 6.2.2.3-1 wiedergegebenen unterschiedlichen Fälle ist zu fragen, welcher Form der Lösungskurven (Stern, Sattelpunkt, Knoten 1. Art etc.) die empirisch ermittelten Daten am nächsten kommen. Das Problem der Analyse eines Regelungsprozesses wird also hier zunächst heuristisch durch eine qualitative Interpretation angegangen, indem die Kurvenverläufe der Herz- und Atemfrequenzwerte der Versuchspersonen mit Hilfe von Methoden der interaktiven graphischen Datenverarbeitung auf einem Bildschirm aufgezeichnet werden.

In einem aufbauenden, allerdings weitreichendem Schritt und unter Aufnahme theoretischer Grundlagen der Psychophysiologie ließe sich neben richtungsweisenden Hypothesen auch ein DGL-System erstellen, das an den vorliegenden Daten überprüft werden könnte (6/12). Da die bisher erzielten Kenntnisse über intrasubjektiv realisierte Strategien zur Regulierung der Herz- und Atemfrequenz in Abhängigkeit zu psychologischen Parametern jedoch noch zu wenig abgesichert und unzureichend sind, wird im folgenden Punkt die Darstellung der sich in der Phasenebene ergebenden Kurvenverläufe beschrieben und es wird ansatzweise analysiert, in welchem Ausmaß aus dem Datenmaterial Hinweise über gelungene Regelungsstrategien abzuleiten sind. Im Idealfall ergeben sich Kurven, die als ausreichende Näherung zu den beschriebenen Zuständen der Stabilität eines Systems betrachtet werden können.

6.3.2 DARSTELLUNG UND DISKUSSION DER EMPIRISCHEN DATEN

Bei der visuellen Beurteilung der auf einem Bildschirm interaktiv ausgegebenen Lösungskurven für die Momentanherzfrequenz und die Werte der Atemfrequenz ergibt sich zunächst eine sehr uneinheit-

liche Struktur. Nachfolgend sind beispielhaft einige der insgesamt 6x41 auf dem Bildschirm erzeugten Kurven wiedergegeben. Die Abszisse markiert das relevante Beobachtungsintervall für die Herzfrequenzwerte, während die Ordinatenachse die bereits bei der Datenregistrierung in ein definiertes Intervall abgebildeten Werte der Atemtätigkeit kennzeichnet.

Versucht man die einzelnen Kurven in den Abbildungen 6.3.2-1/12 anhand von standardisierbaren Kriterien auf visueller Basis bestimmten Kategorien zuzuordnen, so fällt bei oberflächlicher Betrachtung auf, daß in Beziehung zu den bisherigen theoretischen Ausführungen die empirischen Lösungskurven wenn auch grobe, so doch registrierbare Parallelen zu den in der Abb. 6.2.2.3-1 aufgezeich-

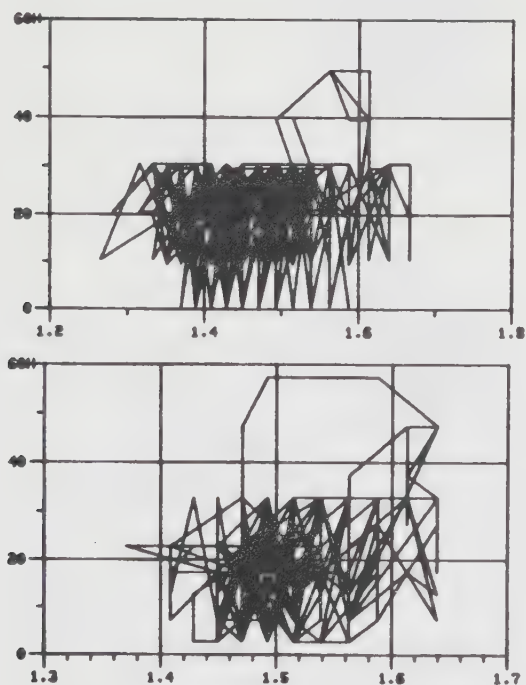


Abb. 6.3.2-1 EMPIRISCH ERMITTELTE LÖSUNGSKURVEN

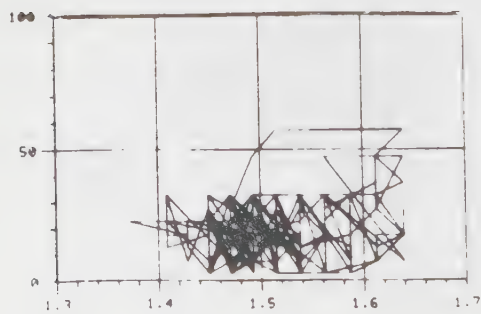
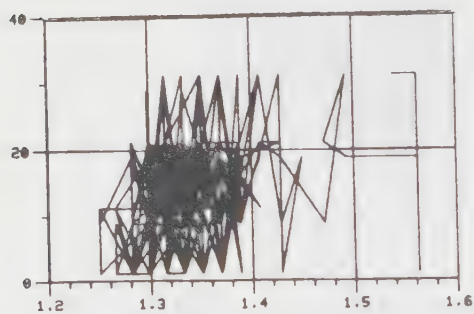
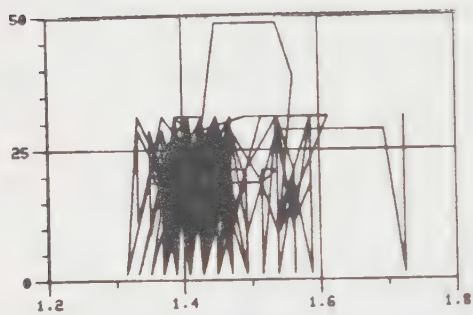
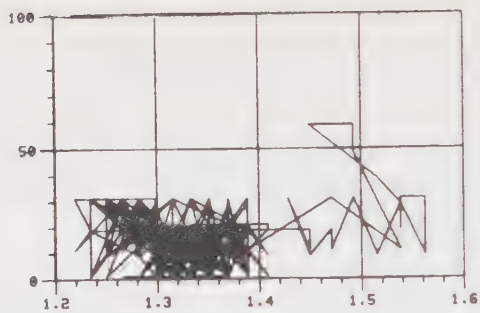


Abb.
6.3.2-1

neten Zuständen des Gleichgewichts oder der (asymptotischen) Stabilität eines Systems in der Form aufweisen, daß für den Untersuchenden kennzeichnende Zusammenballungen oder Agglomerationen der Daten auftauchen, ohne daß diese jedoch als hinreichend exakte Näherung zu einem speziellen Fall aus (6.2.2.3) gewertet werden können. Die im Dialogverkehr bzw. während der Erzeugung des Bildes sichtbare Abfolge der Werte, die aus der hier möglichen Darstellungsform nicht mehr dechiffriert werden kann, läßt Gemeinsamkeiten zwischen den einzelnen Verläufen dahingehend erkennen, daß zu Beginn der Trainingsphasen die Daten von relativ extremen Positionen ausgehen und mit zunehmender Versuchsdauer immer mehr in einem bestimmten, in der Kopie stark dunkel straffierten Sektor zusammenlaufen. Obwohl sich weder intra- noch interindividuell konstante Regulationsmuster oder Abläufe ergeben, wird deutlich, daß mittels der zuvor eingeführten formalen Begriffe und der daraus entwickelten Modellvorstellung diese ansonsten einer Interpretation bzw. Auswertung nicht zugänglichen Befunde in ein Bezugssystem gestellt werden können. Bei einer entsprechenden visuellen Auswertung ist im übrigen zu beachten, daß mitunter ein anderer, zur graphischen Ausgabe geeigneterer Abszissenmaßstab gewählt worden ist und ebenso die Ordinateneinheiten nicht konstant sein müssen.

Zu fragen bleibt, in welchem situativen Kontext ein sich auf einen abgrenzbaren Bezirk richtender Regulationsvorgang funktional ist. Inwieweit also eine solche Konvergenz der Daten die Annahme stützt, daß unter differenten Feedback-Bedingungen im Verlaufe des Versuchs ein Zustand angestrebt und partiell erreicht wird, der physiologisch im Sinne eines relativen Gleichgewichts oder der Entspannung begriffen werden kann und sich formal in einer Agglomeration mittelbar manifestiert, kann jedoch gemessen an methodischen Kriterien aus dem vorliegenden Datenmaterial nicht beantwortet werden. Würde man ein konstantes Reaktionsmuster postulieren und nachweisen wollen, ist dies sicherlich verfehlt, da die bestehenden intra- und interindividuellen Analogien auf visueller Grundlage nicht eindeutig genug kategorisiert werden können.

Generell ist festzuhalten, daß entgegen der eingangs formulierten Erwartung nur sehr vorsichtig psychophysiologische Hinweise aus dem vorliegenden Datenmaterial gewonnen werden können. Ausgehend von den definierten Zuständen der (asymptotischen) Stabilität eines Systems und ihrer graphischen Darstellung für den vereinfachten zweidimensionalen Fall zeigen sich in den empirischen Lösungskurven charakteristische Agglomerationen oder Zusammenballungen, die darauf hindeuten, daß die zu Beginn einer Versuchsphase registrierbaren Abweichungen reduziert werden und der Regulationsprozeß sich auf einen bestimmten, relativ konstanten Zustand einpendelt.

Die Relativität und Vorläufigkeit der Ergebnisse dürfte jedoch nicht prinzipiell gegen die Anwendung eines solchen Lösungsansatzes sprechen als vielmehr darauf aufmerksam machen, daß der Komplexitätsgrad des zu untersuchenden psychophysiologischen Prozesses und des benutzten Modells divergieren und von dort differenziertere, allerdings primär im mathematischen Bezugssystem justierte Modelle mit größerer Wahrscheinlichkeit zu den gesuchten Aussagen führen.

Elemente für alternative Methodenansätze der Datenanalyse bieten sich aus mehreren Teilgebieten der Mathematik bzw. ihrer Nachbardisziplinen wie etwa der Differentialgleichungstheorie, Approximationstheorie, statistischen Übertragungstheorie, Graphentheorie oder der stärker anwendungsorientierten Informationstheorie, biologischen Kybernetik und Mustererkennung an. Inwieweit der Einsatz entsprechend ausformulierter Modelle in der Psychophysiologie über wissenschaftsimmanente Lösungen hinaus zu einem Fortschritt im klinisch-humanen Sinne beizutragen vermag, hängt neben einer Reihe anderer wichtiger Faktoren auch davon ab, in welchem Ausmaß die in dieser Ausarbeitung nur bedingt aufzeigbare gegenstandskonstitutive Funktion formaler Prozesse reflektiert und die Bedingtheit der erzielten und psychophysiologisch als relevant bewerteten inhaltlichen Resultate von den Implikationen und Prämissen des benutzten Modells in Rechnung gestellt wird.

7 ZUSAMMENFASSUNG

Diese Arbeit hat zum Ziel, Ansätze und Verfahren zur Analyse sowie zur numerischen Klassifikation psychophysiologischer Signale nach präzisierten Kriterien aufzugreifen, einführend zu beschreiben und anhand eines speziellen Datensatzes aus dem Bereich der Biofeedback-Forschung ansatzweise zu diskutieren.

So wird im ersten einleitenden Punkt nach Angabe der tangierten Ebenen psychophysiologischer Analyse und nach Formulierung der sich im Kontext der Biofeedback-Forschung stellenden Zielfragen die zugrundeliegende Problemstellung begründet und ausführlicher erörtert. Aus diesen Darlegungen ist der Aufbau dieser primär aus methodischer Sicht konzipierten Arbeit abzuleiten, der aus fünf aufeinanderbezogenen Hauptpunkten besteht.

Zunächst werden wichtige Ansätze und Methoden der Darstellung psychophysiologischer Signale im Zeitbereich behandelt. Einen breiteren Raum nehmen Fragen und Probleme der adäquaten Parameterextraktion ein.

Anschließend werden relevante Strategien der Frequenzbereichsdarstellung von Signalen skizziert, was die Erläuterung tragender Begriffe und Definitionen wie Convolution und Spektrum einschließt.

In einem gesonderten Punkt werden Verfahren und Modelle der numerischen Klassifikation erläutert.

Die bisherigen Ausführungen werden im nächsten Punkt auf einen speziellen empirischen Datensatz bezogen, indem ausgewählte Verfahren der numerischen Klassifikation auf die in den Spektralbereich transformierten Daten angewandt werden. Die erzielten, in dieser Form bislang nicht nachgewiesenen Resultate werden hinsichtlich ihrer psychophysiologischen Bedeutung ansatzweise diskutiert.

Ein alternativer und auf Begriffen aus der mathematischen Theorie der Differentialgleichungen basierender Ansatz ist Gegenstand der folgenden Betrachtung. Es zeigt sich, daß die empirischen Daten mit angebbaren Einschränkungen dem Zustand der asymptotischen Stabilität eines Systems entsprechen.

- 8 ANHANG
- 8.1 PROGRAMME
- 8.1.1 PROGRAMME ZUR DATENORGANISATION

```

PPP:  PROC OPTIONS(MAIN);
      /* PROGRAMM PPP   ZUR DATENORGANISATION
        PUNKT 5.4.1
        ARTEFAKTSELEKTION UND
        NEUE RECORDSTRUKTUR                                */

DCL
  K1 CHAR(80) BASED(P),
  (ALT#,NEU#) CHAR(3),
  (I,Z1,Z2,ZEIT) BIN FIXED(31),
  RTIME EXTERNAL ENTRY,
  (KALT,KT(20)) CHAR(80),
  PH CHAR(900) VAR,
  ALTER_WERT CHAR(3) INIT('000'),
  #DATEN1 CHAR(2) INIT(' '),
  #DATEN2 CHAR(2) INIT(' '),
  (NSKIP,NN) BIN FIXED(31),
  FEHLER(5) BIN FIXED;

DCL
  1 KALT1 DEFINED KALT POS(1),
  2 WERT(10),
  3 VC CHAR(2),
  3 ATEM CHAR(2),
  3 HERZ CHAR(3),
  2 IDENT CHAR(10),
  AC CHAR(2);
  /* RECORD STRUKTUR */

DCL
  1 REC_DATEN1 UNALIGNED,
  2 LFDNR CHAR(4),
  2 VAPBN CHAR(4),
  2 #ARTEFAKTE BIN FIXED,
  2 IV(60) BIN FIXED,
  2 #PHASEN BIN FIXED,
  2 PHASEN(60),
  3 PH_ART CHAR(2),
  3 TRAINING JA CHAR(2),
  3 PH_ENDE BIN FIXED,
  WERTE_VERCODED CHAR(12000) VAR;
  /* SUBROUTINE VERCODEN */

DCL
  XI(10) BIN FIXED,
  XJ(10) BIN FIXED,
  LJ CHAR(20) DEFINED XJ,
  LI CHAR(20) DEFINED XI;
  /* PARAMETER SETZEN */
  LL IP=0;
  IP,I=0; FEHLER=0;
  DO NN=1 TO 60;
  TRAINING JA(NN)='97';
  END;

```

```

ON ENDFILE (BAND) GOTO ENDE;
/* RESTART => ANFANGSWERT */
GET LIST (NSKIP);
PUT SKIP LIST('ERSTE KARTE =', NSKIP);
LESEN:
  READ FILE(BAND) SET(P);
  I=I+1;
  IF I< NSKIP THEN GOTO LESEN;
  ON ERROR PUT DATA (I,IP,KT);
  ON CONVERSION BEGIN;
  DCL (ONSOURCE,ONCHAR) BUILTIN;
  FEHLER(3)=FEHLER(3)+1;
  PUT EDIT('----> ',ONCHAR){SKIP,A,A};
  ONCHAR='0';
  END;
  CALL RTIME(Z1);
  READ FILE (BAND) SET(P); I=I+1;
NEUER_PROBAND:
  IP=IP+1;
  WERTE_VERCODED= #DATEN1 || #DATEN2 ;
KEIN_N_P:
  DO L=1 TO 20 WHILE(SUBSTR(K1,79,1)~='V');
  KT(L)=K1;
  READ FILE (BAND) SET(P); I=I+1;
  END;
  /* BEI MEHR ALS 20 VORLAUFKARTEN */
  IF L>20 THEN DO;
  FEHLER(1)=FEHLER(1)+1;
  PUT EDIT('&&&>> L>20 ')(SKIP,A);
  GOTO KEIN_N_P;
  END;
  NEU#,ALT#=SUBSTR(K1,76,3);
  PUT EDIT(L-1,' VORLAUFKARTEN ---> ',
  SUBSTR(K1,72,9)) (SKIP,F(6).2A);
  J,LV=1;
  DO LL=1 TO L-1;
  /* ANNAHME - NUMERISCHE ZEICHEN */
  IF VERIFY(KT(LL),' ')~=0
  & SUBSTR(KT(LL),5,8)~='17181718'
  & INDEX(SUBSTR(KT(LL),41,3),' 14')=0
  & SUBSTR(KT(LL),1,7)~='248817 '
  & SUBSTR(KT(LL),1,2)~='//'
  & VERIFY(KT(LL),' 0123456789')=0 THEN DO;
  /* FESTSTELLEN, WELCHE VERB.NR */
  DO JX=LV TO LV+19
  WHILE(IV(JX)~=0 & IV(JX)~=7777 & IV(JX)~=8888);
  GET STRING(KT(LL)) EDIT
  ((IV(J) DO J=LV TO LV+19))(F(4));
  END; J=JX;
  LV=JX;
  END;
  END;
  LV=J-1;
  IF LV >80 THEN FEHLER(2)=FEHLER(2)+1;
  PUT EDIT(' VERB.-->',LV,(IV(J) DO J=1 TO LV))
  (A,F(4),X(2),(LV) F(5));
  LVV=LV;
  LK=0;
  AC=SUBSTR(K1,1,2); PH=' ';
  LV=0; IPHAS=0;
  JO,JJO=1;
  PUT SKIP;
  DO WHILE(SUBSTR(K1,79,1)='V' & ALT#=#NEU#);

```

```

LK=LK+1;
KALT=K1;
/* ARTEFAKT VERB. */
DO J=JO TO LVV BY 1
WHILE(LK=CEIL(IV(J)/BINARY(10)));
IWO=MOD(IV(J),10);
IF IWO=0 THEN IWO=10;
IDORT=(IWO-1)*7 +5;
IF IWO=1 THEN SUBSTR(KALT,IDORT,3)=ALTER_WERT;
ELSE
SUBSTR(KALT,IDORT,3)=SUBSTR(KALT,IDORT-7,3);
JJO=JJO+1;
END;
JO=JJO;
LA=1;
AA:
DO L1=LA TO 10 WHILE(AC=VC(L1) |VC(L1)='16');
LV=LV+1;
END;
IF L1<11 THEN DO;
IPHAS=IPHAS+1;
PH=PH || AC;
PH_ART(IPHAS)=AC;
PH_ENDE(IPHAS)=LV;
AC=VC(L1);
LA=L1;
GOTO AA;
END;
ALTER_WERT=SUBSTR(KALT,68,3);
CALL VERCODEN;
READ FILE(BAND) SET(P); I=I+1;
NEU#=SUBSTR(K1,76,3);
END;
IPHAS=IPHAS+1;
PH=PH || AC;
PH_ART(IPHAS)=AC;
PH_ENDE(IPHAS)=LV;
PUT EDIT('***> PROBAND ',KALT,' HAT ',LK,' KARTEN')
(SKIP,3 A,F(7),A);
PUT EDIT(PH) (SKIP,A);
PUT EDIT((132)'-' ) (SKIP,A);
CALL OUTREC;
/* RESTZEITABFRAGE (SPEZ.IBM 360/50 RZ MS */
CALL RTIME(Z2);
IF Z2 <= 2000 THEN GOTO FINISH;
GOTO NEUER_PROBAND;
ENDE:
IPHAS=IPHAS+1;
PUT EDIT ('|', IPHAS,AC,LV)
(A(1),F(2),X(1),A(2),F(5));
PH=PH || AC;
PH_ART(IPHAS)=AC;
PH_ENDE(IPHAS)=LV;
CALL VERCODEN;
CALL OUTREC;
LLIP=LLIP+1;
PUT EDIT ('***=> PROBAND ',SUBSTR(KALT,72,9),
' HAT ',LK,' KARTEN') (SKIP,3,A,F(7),A);
FINISH:
PUT EDIT((132)'-' ) (SKIP,A);
CALL RTIME(Z2);
PUT SKIP(1) DATA(FEHLER);
Z2=Z2-2;

```

```

PUT SKIP LIST('ZEIT =',ZEIT,' *10**(-2) SEC. ');
PUT SKIP DATA(I,IP,LLIP);
OUTREC: PROC;
PUT STRING(LFDNR) EDIT(IP)(F(4));
VAPBN=SUBSTR(KALT,76,4);
#ARTEFAKTE=LVV;
#PHASEN=IPHAS;
LK=LK*10;
UNSPEC(#DATEN1)=UNSPEC(LK);
LK=LK+LK;
UNSPEC(#DATEN2)=UNSPEC(LK);
SUBSTR(WERTE_VERCODED,1,4)=#DATEN1 ||#DATEN2;
DO NN=LVV+1 TO 60; IV(NN)=9; FND;
DO NN=#PHASEN+1 TO 60;
PH_ART(NN)='98';
PH_ENDE(NN)=9999;
END;
LLIP=LLIP+1;
WRITE FILE(OUTPUT) FROM(REC_DATEN1);
WRITE FILE(OUTPUT) FROM (WERTE_VERCODED);
END; /* UOUTREC */
VERCODEN: PROC;
DO N=1 TO 10;
GET STRING (ATEM(N)) EDIT (XI(N))(F(2));
GET STRING (HERZ(N)) EDIT (XJ(N))(F(3));
III=2*N;
WERTE_VERCODED=WERTE_VERCODED || SUBSTR(L1,III,1)
|| SUBSTR(LJ,III,1);
END;
END; /* VERCODEN */
END; /* PPP-MAIN*/

```

```

ABC:  PROC OPTIONS(MAIN);
      /* PROGRAMM ABC
          PUNKT 5.4.1
          SORTIEREN UND WEGSCHREIBEN
          NACH VERSUCHSGRUPPEN (A,B,C)  */

DCL
1 REC_DATEN1 UNALIGNED,
2 LFDNR CHAR(4),
2 VAPBN CHAR(4),
2 #ARTEFAKTE BIN FIXED,
2 IV(60) BIN FIXED /* ARTEFAKT_NR */ ,
2 #PHASEN BIN FIXED,
2 PHASEN(60),
  3 PH_ART CHAR(2),
  3 TRAINING JA CHAR(2),
  3 PH_ENDE BIN FIXED,
WERTE_VERCODED CHAR(12000) VAR;

DCL
VERGLEICH(41) CHAR(7),
ALT_RECORD BIN FIXED,
IREC BIN FIXED;
IREC=0;
GET EDIT(ALT_RECORD) (COL(1).F(3)) COPY;
GET EDIT (VERGLEICH) (COL(1).A(7)) COPY;
ON ENDFILE(BAND) GOTO ENDE;

LESEN:
  READ FILE(BAND) INTO (REC_DATEN1);
  READ FILE (BAND) INTO (WERTE_VERCODED);
  IREC=IREC+2;
  IF IREC < ALT_RECORD THEN GOTO LESEN;
  DO I=1 TO 41;
    IF VAPBN = SUBSTR(VERGLEICH(I),1,4)
      THEN GOTO WEG;
  END;
  GOTO LESEN;

WEG:
  IF SUBSTR(VERGLEICH(I),7,1) ='A' THEN DO;
    WRITE FILE(RAUSA) FROM (REC_DATEN1);
    WRITE FILE(RAUSA) FROM (WERTE_VERCODED);
    PUT SKIP LIST('VAPBN - GRUPPE A ',VERGLEICH(I));
    GOTO LESEN;
  END;
  IF SUBSTR(VERGLEICH(I),7,1) ='B' THEN DO;
    WRITE FILE(RAUSB) FROM (REC_DATEN1);
    WRITE FILE(RAUSB) FROM (WERTE_VERCODED);
    PUT SKIP LIST('VAPBN - GRUPPE B ',VERGLEICH(I));
    GOTO LESEN;
  END;
  IF SUBSTR(VERGLEICH(I),7,1) ='C' THEN DO;
    WRITE FILE(RAUSC) FROM (REC_DATEN1);
    WRITE FILE(RAUSC) FROM (WERTE_VERCODED);
    PUT SKIP LIST('VAPBN - GRUPPE C ',VERGLEICH(I));
    GOTO LESEN;
  END;
  GOTO LESEN;
ENDE: END;

```

8.1.2 PROGRAMME ZUR FREQUENZBEREICHSDARSTELLUNG

```

C      PROGRAMM AMS      ZUR DATENDESCRIPTION
C                          PUNKT 5.4.2
C                          BERECHNUNG DER AMPLITUDENSPEKTREN
C                          DER MOMENTANHERZFREQUENZ
      DIMENSION T(1000),Y(1000),ETA(1000),A(1000),B(1000)
      INTEGER*2 N1,N2,NPHAS,PHA(60),PHEN(60)
      1      IPH1/'17'/,IPH2/' 1'/,IPH3/' 1'/
      2      WERTE(12000)
      EQUIVALENCE (WERTE(1),A(1)),(WERTE(4001),B(1)),
      1      (WERTE(8001),ETA(1))
      COMMON WERTE
      IOUT=6
      DO 1 I=1,12000
      1  WERTE(I)=0
      INT=300
      RESTART => IANF=1.PBN
      READ(5,2000) IANF
      WRITE(6,2001) IANF,INT
      WRITE(6,2002) IPH1,IPH3
      DO 10 IJ=1,400
      CALL RTIME(III)
      RESTZEITABFRAGE (SPEZ. IBM 360/50 RZ MS)
      IF(III.LE.6900) GOTO 999
      CALL DATA (NR,N1,N2,NPHAS,PHA,PHEN,IER,IANF)
      IF (IER.F0.991) GOTO 999
      WRITE(8) WERTE
      NP=NPHAS-1
      DO 10 I=1,NP
      IPP=I
      REWIND 8
      READ(8) WERTE
      L=(PHEN(I)*2)+1
      LO=L
      NZ=PHEN(I+1)-PHEN(I)
      IF(NZ.LT.INT) GOTO 40
      IF(PHA(IPP+1).NE.IPH1.AND.PHA(IPP+1).NE.IPH2)
      1      GOTO 40
      WRITE(6,1000) NR,IPP,PHEN(IPP),PHEN(IPP+1),
      1      NZ,LO,PHA(IPP+1),NP
      DO 30 J=1,NZ
      Y(J) =WERTE(L+1)
      L=L+2
      30 CONTINUE
      CALL AMPSP (Y,NZ,T,ETA,A,B,NR,PHEN,PHA,IPP)
      GOTO 10
      40 WRITE(6,1001) NR,IPP,PHEN(IPP),PHEN(IPP+1),
      1      NZ,LO,PHA(IPP+1),NP
      10 CONTINUE
      999 WRITE(IOUT,9000) IER
      STOP
      1000 FORMAT(9X,'***-NR=',A4,' BEARB.PHASE  =',I3,' INTERVALL< >=',
      1      2I5,' NZ=',I5,' L(=2*)=',I5,' PHA(IPP+1)=',A2,
      2      ' NP=',I5)
      1001 FORMAT(9X,'PHN-NR=',A4,' AUSGEE.PHASE  =',I3,' INTERVALL< >=',
      1      2I5,' L(=2*)=',I5,' PHA(IPP+1)=',A2,

```

```

2000 FORMAT(I3)
2001 FORMAT(9X,'PROGRAMM AMP',/9X,'1.PBN-NR (=IANF) = ',I3,
1      /9X,'VORGEgebenES MINIMAL INTERVALL (=INT) = ',I5)
2002 FORMAT(9X,'VERSUCHSGRUPPE A - => IPH1=',A2,
1      /9X,'VERSUCHSGRUPPE B - => IPH2=',A2,
2      /9X,'VERSUCHSGRUPPE C - => IPH3=',A2)
9000 FORMAT(9X,'ENDE - MAIN / ZEIT - DATEN || IER=',I5)
END

```

```

C      SUBROUTINE DATA (NR,N1,N2,NPHAS,PHA,PHEN,IER,IANF)
      EINLESEN DER DATEN - BANDVERS.
      LOGICAL*1 LW(2,12000)
      INTEGER*2 WERTE(12000)
      LOGICAL*4 VAPBN
      INTEGER*2 N1,N2,NARF,IV(60),NPHAS,PHA(60),PHEN(60),TRJ(60)
      INTEGER*2 IREC/0/
      EQUIVALENCE (LW(1,1),WERTE(1))
      COMMON WERTE
      IIN=15
10     READ(IIN,END=20) NR,VAPBN,NARF,IV,NPHAS,
1      (PHA(I),TRJ(I),PHEN(I),I=1,60)
      READ(IIN,END=20) N1,N2,(LW(2,I),I=1,N2)
      IREC=IREC+2
      INR=IREC/2
      WRITE(6,2000) INR,NR,VAPBN,IREC
C      RESTART - FORTLAUFENDE ZAEHLUNG
      IF(INR.LT.IANF) GOTO 10
      IER=0
      RETURN
20     IER=991
      RETURN
2000 FORMAT(9X,'EINGELESEN INR=',I5,' NR=',A4,' VAPBN ',A4,2X,I5)
END

```

```

C      SUBROUTINE AMPSP (Y,N, T,ETA,A,B,NR,PHEN,PHA,IPP)
      BERECHNUNG DER AMPLITUDENSPEKTREN
      DIMENSION T(1000),Y(1000),ETA(1000),A(1000),B(1000)
      INTEGER*2 N1,N2,NPHAS,PHA(60),PHEN(60)
      INTEGER NRA/500/
      IF(NRA.NE.NR) KK=0
      NRA=NR
      KK=KK+1
      T(1)=Y(1)
      DO 9 I=2,N
      T(I)=T(I-1)+Y(I)
9     CONTINUE
      WRITE(6,1003)
      WRITE(6,1001) T(1),T(N)
      DO 10 I=1,N
      Y(I)=100 ./Y(I)
10    CONTINUE
      WRITE(6,1002) (Y(I),I=1,10)
      WRITE(6,1002) (Y(I),I=297,N)
      TMAX=100./C
      TMIN=0./C
      DO 11 I=1,N

```

```

      IF (T(I).LT.TMIN) TMIN=T(I)
      IF (T(I).GT.TMAX) TMAX=T(I)
11  CONTINUE
      TFAK=6.283/(TMAX-TMIN)
      DO 12 I=1,N
12  T(I)=(T(I)-TMIN)*TFAK
      WRITE(6,1004) TFAK,TMIN,TMAX
      WRITE(6,3001) T(1),T(N)
C  DATENPUNKTE AEQUIDISTANT
      DELTA=6.283/FLOAT(N)
      K=1
      DO 20 I=1,N
      TAU1=FLOAT(I)*DELTA
      DO 25 L=K,N
      IF (T(L).GT.TAU1) GOTO 30
25  CONTINUE
30  CONTINUE
      L=L-1
      K=L
      ETA(I)=(Y(L+1)-Y(L))/(T(L+1)-T(L)) * (TAU1-T(L)) + Y(L)
20  CONTINUE
      WRITE(6,2000)
      WRITE(6,2001) (T(I),ETA(I),Y(I),I=1,N)
      NM=N/2
      M=NM
      WRITE(6,2002) N,NM,M
      CALL FKOEFF (ETA,NM,M,A,B,IER)
      WRITE(6,2500) IER
      WRITE(6,3000)
      WRITE(6,3001) (A(I),B(I),I=1,M)
      XN=NM
      T(1)=0.0
      T(2)=1.0/XN
      DO 70 I=3,M
70  T(I)=T(I-1)+T(2)
      DO 80 I=2,M
80  A(I)=(A(I)**2+B(I)**2)**0.5
      CALL SAVE (T,A,M,IPP,PHEN,PHA,NR,KK)
      CALL ZEICH(T,A,M,IPP,PHEN,PHA,NR,KK)
      RETURN
1000 FORMAT(16(2X,F3.0))
1002 FORMAT(9X,10F6.2,'|',10F6.2)
1003 FORMAT(9X,'X-ACHSE(ZEIT T) KUMMULIERT - T(1) | T(N)')
1004 FORMAT(9X,'TFAK=',F16.4,' TMIN=',F16.4,' TMAX=',F16.4)
2000 FORMAT(9X,'T(I)          ETA(I)          Y(I)')
2001 FORMAT(9X,12F10.3)
2002 FORMAT(9X,'N=',I5,' NM=',I5,' M=',I5)
2500 FORMAT(9X,'IER=',I5)
3000 FORMAT(9X,'COS - SIN KOEFFIZIENTEN')
3001 FORMAT(9X,10F12.4)
4000 FORMAT(9X,'ARTEF.NR.=',I3,' Y(I)= ',F10.5)
      END

SURROUTINE FKOEFF (FNT,N,M,A,B,IER)
C  BERECHNUNG DER COS/SIN-KOEFF.
C  SSP-MANUAL VERS. III, P.275
      DIMENSION A(100),B(100),F(100)
      IER=0
20  IF(M) 30,40,40
30  IF=2

```

```

RETURN
40 IF (M-N) 60,60,50
50 IER=1
RETURN
60 AN=N
COEF=2.0/(2.0*AN+1.0)
CONST=3.141593*COEF
S1=SIN(CONST)
C1=COS(CONST)
C=1.0
S=0.0
J=1
FNTZ=FNT(1)
70 U2=0.0
U1=0.0
I=2*N+1
75 U0=FNT(I)+2.0*C*U1-U2
U2=U1
U1=U0
I=I-1
IF(I-1) 80,80,75
80 A(J)=COEF*(FNTZ+C*U1-U2)
B(J)=COEF*S*U1
IF(J-(M+1)) 90,100,100
90 Q=C1*C-S1*S
S=C1*S+S1*C
C=Q
J=J+1
GOTO 70
100 A(1)=A(1)*0.5
RETURN
END

```

```

SUBROUTINE SAVE (FR,AM,M,IPP,PHEN,PHA,NR,KK)
C   SPEICHERN DER AMPL./FREQ. PRO PHASE/PBN
DIMENSION FR(1000),AM(1000)
INTEGER*2 PHEN(60),PHA(60)
C   KONTROLLAUSGABE DRUCKER 6
IOUT=6
WRITE(IOUT,6000) NR,KK,M,IPP,PHEN(IPP),PHEN(IPP+1),PHA(IPP+1)
6000 FORMAT(1X,'PBN-NR=',A4,' KK-PHASE=',I5,' N-PUNKTE.FREQ.=',I5,
1      ' IPP-URSP.PHASE/PP2=',I5,' INTERVALL < > =',2I5,
2      ' PHASEN-KENNUNG = ',A2)
WRITE(IOUT,6001) (FR(I),AM(I),I=1,M)
6001 FORMAT(9X,8F10.4)
C   WEGSCHREIBEN AUF BAND 1
IOUT=1
WRITE(IOUT,100) NR,KK,M,IPP,PHEN(IPP),PHEN(IPP+1),PHA(IPP+1)
WRITE(IOUT,101) (FR(I),AM(I),I=1,M)
100 FORMAT(A4,1X,5I5,A2)
101 FORMAT(8F10.5)
RETURN
END

```

```

SUBROUTINE ZEICH (X,Y,N,IPP,PHEN,PHA,NR,KK)
ZEICHNEN BZW. PLOTTEN
DRUCKER-VERSION
DIMENSION X(1000),Y(1000),A(1000)
INTEGER ZEILE(100)/100*' '/,STERN/'*'/,BLANK/' '/,PUNKT/'.'/,
*      DRUP0,SEITE
  INTEGER*2 PHEN(60),PHA(60)
  DO 9 I=1,N
    A(I)=Y(I)
9  CONTINUE
  YQ=Y(1)
  Y(1)=0.0
  YMAX=0.0
  YMIN=1000.0
  DO 10 I=1,N
    IF(Y(I).LT.YMIN) YMIN=Y(I)
    IF(Y(I).GT.YMAX) YMAX=Y(I)
10 CONTINUE
  WRITE(6,1000)
  Y - DRUCKPOSITIONEN 130-31
  YFAK=(130-31)/(YMAX-YMIN)
  WRITE(6,1001) YMIN,YMAX,YFAK
  DO 20 I=1,N
    Y(I)= (Y(I)-YMIN)*YFAK
20 CONTINUE
  WRITE(6,1010)
  WRITE(6,1009) NR,KK,N,IPP,PHEN(IPP-1),PHEN(IPP),PHA(IPP)
  WRITE(6,1006) X(1),YQ
  DO 50 I=2,N
    DRUP0=1.5+ Y(I)
    ZEILE(DRUP0)=PUNKT
    WRITE(6,1006) X(I),A(I),(ZEILE(K),K=1,DRUP0)
    ZEILE(DRUP0)=BLANK
50 CONTINUE
  RETURN
1000 FORMAT(9X,'DATEN-BILD | PARAMETER')
1001 FORMAT(9X,'YMIN=',F12.3,' YMAX=',F12.3,' YFAK=',F12.3)
1006 FORMAT(1X,'X=',F8.4,1X,'Y=',F8.3,T31,100A1)
1009 FORMAT(1X,'PBN-NR=',A4,' KK-PHASE=',I5,' N-PUNKTE.FRQ.=',I5,
1      ' IPP-URSP.PHASE/PP2=',I5,' INTERVALL < > =',I5.
2      ' PHASEN-KENNUNG = ',A2)
1010 FORMAT('')
END

```

8.1.3 PROGRAMME ZUR GRAPHISCHEN AUSGABE ("HIDDEN-LINE")

```

C      PROGRAMM HID   ZUR DATENDESKRIPTION
C      PUNKT 5.4.2
C      ('HIDDEN-LINE' TECHNIK)
C
      DIMENSION D(1000),HL(5000)
      IIN=3
10     CONTINUE
      READ(IIN,1000,END=99)      N,M,NK,BL,BB,
1                                   XFAK,YFAK,ZFAK,
2                                   ALPHA,ZMIN,ZMAX
      WRITE(6,2000)              N,M,NK,BL,BB,
1                                   XFAK,YFAK,ZFAK,
2                                   ALPHA,ZMIN,ZMAX
      NU6=6
      CALL PLOTS
      WRITE(6,200) IER
      CALL MTMAP(BL,BB,IER)
      CALL SPROJ(N,M,ALPHA,ZMIN,ZMAX,BL,BB,XFAK,YFAK,
1              ZFAK,D,HL,NK,6)
      CALL PLOTE
      CALL PLOTS
      GO TO 10
99     CALL PLUTE
      CALL PLOTPV
      STOP
200    FORMAT(20X,'IER=',I4)
1000   FORMAT(314.8F7.3)
2000   FORMAT(1X,'PARAMETER :',3I5,4X,8F7.3)
      END

```

```

C      SUBROUTINE MTMAP(BL,BB,IER)
C      MEHRERE BILDER AUF EINE SEITE
      REAL RX/0./,RY/0./,XMAX/0./,BBL/0./
      IER=0
      IF(BB.LE.68.) GO TO 10
      IER=1
      RETURN
10     CONTINUE
      IF(RY-BBL+BB.LE.68.) GO TO 20
      CALL PLOT(XMAX,-RY,-3)
      RX=RX+XMAX
      RY=0.
      XMAX=0.
      GO TO 30
20     CONTINUE
      CALL PLOT(0.,BBL,-3)
      RY=RY+BBL
30     CONTINUE
      BBL=B-
      IF(XMAX.LT.BL) XMAX=BL
      IF(RY-BBL.LE.100.) RETURN
      IER=2

```

```

SUBROUTINE SPROJ(N,M,ALPH,ZMIN,ZMAX,BL,BB,XFAK,YFAK,
1          ZFAK,D,HL,NK,NU6)
DIMENSION D(N),HL(NK)
REAL ARG/1.745329E-02/
RU=3.
RL=3.0
IF(XFAK.LE.0.) XFAK=1.
IF(YFAK.LE.0.) YFAK=1.
IF(ZFAK.LE.0.) ZFAK=1.
SAL=SIN(ARG*ALPH)*YFAK
CAL=COS(ARG*ALPH)*YFAK
N1=N-1
M1=M-1
DZ=ZMAX-ZMIN
XF=(BL-RL)/(N1*XFAK+M1*CAL)
YF=(BB-RU)/(M1*SAL+DZ*ZFAK)
DX=XF*XFAK
DXY=XF*CAL
DY=YF*SAL
ZF=YF*ZFAK
YIN=CAL/XFAK
NHL=N+M1*CAL/XFAK+0.8
WRITE(NU6,1000) NK,NHL
1000 *  FORMAT(1X,10(' '), '  DER VEKTOR HL HAT DIE LAENGE ',I4/
      23X,'DIE MINDESTLAENGE BETRAEGT ',I4,3X,10(' '))
CALL PLOT(RL,RU,-3)
CALL FACTOR(0.94)
CALL AXIS(0.,0., 80.,0.,0.,0.05,'FREQUENZ',8)
CALL AXIS(0.,0.,25.0,90.,0.,0.10,'AMPLITUDE',-9)
CALL PLOT(N1*DX,0.,2)
CALL PLOT(0.,0.,3)
CALL PLOT(M1*DXY,M1*DY,2)
CALL PLOT(0.,0.,3)
CALL PLOT(0.,ZF*DZ,2)
CALL PLOT(0.,0.,3)
K=1
CALL DATMAT(D,N,K,KENN)
CALL SYMBOL(0.,0.,0.1,PKENN,0.,8)
Y=ZF*(D(1)-ZMIN)
HL(1)=Y
CALL PLOT(0.,Y,2)
X=0.
DO 10 I=2,N
Y=ZF*(D(I)-ZMIN)
X=X+DX
CALL PLOT(X,Y,1)
10  HL(I)=Y
CONTINUE
N1=N+1
DO 20 I=N1,NHL
20  HL(I)=0.
CONTINUE
XV=0.
YV=0.
DO 60 K=2,M
XV=XV+DXY
YV=YV+DY

```

```

      IV=(K-1)*YIN + 0.5
      CALL DATMAT(D,N,K,KENN)
      Y=ZF*(D(1)-ZMIN) + YV
      X=XV
      IMOVE=3
      CALL PLOT(X,Y,IMOVE)
      IIV=1+IV
      IF(HL(IIV).GT.Y) GO TO 30
      IMOVE=2
      HL(IIV)=Y
      CALL PLOT(X,Y,IMOVE)
30    CONTINUE
      DO 60 I=2,N
      Y=ZF*(D(I)-ZMIN) + YV
      X=X+DX
      IIV=1+IV
      IF(HL(IIV).GT.Y) GO TO 50
      HL(IIV)=Y
      IF(IMOVE.NE.3) GO TO 40
      IMOVE=2
      CALL PLOT(X,Y,3)
      GO TO 60
40    CONTINUE
      CALL PLOT(X,Y,2)
      GO TO 60
50    CONTINUE
      IMOVE=3
60    CONTINUE
      CALL PLOT(-RL,-RU,-3)
      RETURN
      END

```

```

C      SUBROUTINE DATMAT(D,N,K,PKENN)
      DATENEINGABE
      DIMENSION D(N)
      READ(5,1000) (D(I),I=1,N)
      WRITE(6,2000) (D(I),I=1,N)
      DMIN=1000.0
      DMAX=0.0
      DO 10 I=1,N
      IF(D(I).LT.DMIN) DMIN=D(I)
      IF(D(I).GT.DMAX) DMAX=D(I)
10    CONTINUE
      DFAK=100.0/(DMAX-DMIN)
      DO 20 I=1,N
      D(I)=(D(I)-DMIN)*DFAK
20    RETURN
1000  FORMAT(8F10.5)
2000  FORMAT(1X,10F10.5)
      END

```

8.1.4 PROGRAMME ZUR DATENKLASSIFIKATION

```

C      PROGRAMM KLS      ZUR DATENKLASSIFIKATION
C                          PUNKT 5.4.3
C                          PARAMETER ISTRAT => CLUSTER-
C                          STRATEGIE (A-E)
C
      DIMENSION X(2050),Y(2050),Z(2050),
1      CM(200),CV(200),V(2500),R(1275),EW(50)
      INTEGER NP(41),KCT(50)
      DIMENSION SM(50),SS(50)
      READ(5,1000)          N,M,K1,K2
      WRITE(6,2000)         N,M,K1,K2
      NM=N*M
C
10     READ(5,1001,END=99) ILAUF,ISTRAT
      WRITE(6,2001) ILAUF,ISTRAT
      CALL SELEC (X,Y,N,M,NM,ILAUF)
      WRITE(6,2002) N
      GOTO (1,2,3,1,5),ISTRAT
2     CALL ZTRANS (N,M,X,SM,SS,NM)
1     CALL QANALY (N,M,K1,K2,
1         X,Z,Y,V,R,EW,NP,KCT,CM,CV,
2         NM,M*M,M*(M+1)/2,M*(M-1))
      IF (ISTRAT.EQ.4) GOTO 4
      STOP
3     CONTINUE
      DO 9 I=1,N
9      NP(I)=N
4     CONTINUE
      DO 40 K=K1,K2
      WRITE(6,2003)
      CALL HLCMB1 (N,M,K,X,NP,CM,CV,KCT,SPW,NM,M*K)
      WRITE(6,90) K
      WRITE(6,100) (NP(I),I=1,N)
      WRITE(6,120) (CV(L),L=1,K)
      WRITE(6,110) SPW
      WRITE(6,130)
      DO 30 J=1,K
      IND=(J-1)*M
30     WRITE(6,140) (CM(IND+L),L=1,M)
40     CONTINUE
      STOP
5     CONTINUE
      DO 50 I=1,N
50     NP(I)=0
      IHK=1
      TRESH=1.0
      CALL KTRANS(N,M,M1,IHK,X,Y,V,R,SM,SS,EW,TRESH,NM,M*(M+1)/2,M*M)
      DO 70 K=K1,K2
      WRITE(6,2004)
      MK=M*K
      CALL HLCMB1 (N,M1,K,Y,NP,CM,CV,KCT,SPW,N*M1,M1*K)
      WRITE(6,90) K
      WRITE(6,100) (NP(I),I=1,N)
      CALL MEAS (N,M,K,X,NP,CM,CV,KCT,NM,MK)
      WRITE(6,210)
      DO 60 J=1,M

```

```

60 WRITE(6,140) (CM(L),L=1,MK,M)
70 CONTINUE
C RESTZEITABFRAGE
CALL RTIME(JII)
IF(JII.LE.1000) GOTO 99
GOTO 10
99 WRITE(6,9999)
STOP
90 FORMAT(9X,'CLUSTERLAUF MIT',I3,' CLUSTER'//,
1 9X,'CLUSTERZUGEHÖRIGKEIT:'//)
100 FORMAT(1X,15I3)
110 FORMAT(/9X,'WERT DER ZIELFUNKTION (SPUR W) =' ,F10.3)
120 FORMAT(9X,'CLUSTERVARIANZEN'/(/1X,10(F8.1,1X)))
130 FORMAT(/9X,'GRUPPENCENTROIDE'//)
140 FORMAT(/9X,10F9.3)
210 FORMAT(///9X,'GRUPPENCENTROIDE URSPR.G.DATEN X!')
1000 FORMAT(4I5)
1001 FORMAT(2I2)
2000 FORMAT(9X,'VORGEBENENE PARAMETER '//,
1 9X,'N=',I5,' M=',I5,' K1=',I5,' K2=',I5)
2001 FORMAT('1',8X,'ANALYSIERTE PHASE - ILAUF =' ,I5,
1 9X,'CLUSTER-STRATEGIE =' ,I4//)
2002 FORMAT(9X,'ERMITT. ELEMENTENANZAHL N= ',I5)
2003 FORMAT(///9X,'CLUSTERANALYSE - HILL-CLIMB (HLCMB1)')
2004 FORMAT(///9X,'CLUSTERANALYSE - HILL-CLIMB (HLCMB1)'.
1 /9X,'KTRANS - DATEN')
2005 FORMAT(/9X,'KTRANS - BERECHNETE DIM. M1=' ,I5//)
9999 FORMAT(9X,'E N D E MAIN - ZEITUEBER./KEINE PARAMETERKARTEN')
END

```

```

SUBROUTINE SELEC (X,Y,N,M,NM,ILAUF)
C DATENSELEKTION
DIMENSION X(NM),Y(NM)
IIN=15
IF(ILAUF.EQ.6) N=N-1
NM=N*M
NM2=NM/2
DO 10 I=1,N
READ(IIN,5000,END=99) IG,(X(L),L=1,NM,N)
99 CONTINUE
WRITE(6,7000) ILAUF,I
5000 FORMAT(I4,40A4)
6000 FORMAT(1X,I4,20F6.3)
7000 FORMAT(1X,'ENDBED.99 - ILAUF=' ,I5,' I=' ,I5)
END

```

SUBROUTINE QANALY (N,M,K1,K2,X,Z,Y,V,R,EW,NP,KCT,CM,CV,
 * NM,MM,M2MP1,MK)

Q-ANALYSE ZUR GRUPPIERUNG VON PERSONEN

EINGABE:

N INTEGER ANZAHL DER ELEMENTE
 M INTEGER ANZAHL DER VARIABLEN
 K1,K1 INTEGER BEREICH, INNERHALB DESSEN GRUPPIERUNGEN
 BERECHNET WERDEN SOLLTEN: $K1 \leq K \leq K2$
 X REAL (N,M)-DATENMATRIX
 NM,MM,M2MP1,MK
 INTEGER DIMENSIONS LAENGEN $N \times M, M \times M, M \times (M+1)/2, M \times K$
 AUSGABE:
 NP INTEGER GRUPPIERUNGSVEKTOR WIE IN MINDS1
 CM,CV WIE IN MINDS1
 Z,Y REAL (N,M)-ARBEITSMATRIZEN
 V REAL (M,M)-MATRIX FUEER DIE EIGENVEKTOREN
 R REAL SYMMETRISCHE MATRIX IN 1-DIM SPEICHERUNG
 DER LAENGE $M \times (M+1)/2$ FUEER DIE Q-KORRELATIONSKOEFF.

BENÖTIGTE SUBROUTINEN

EWP ZUR BERECHNUNG DER EIGENWERTE UND EIGENVEKTOREN
 EINER REELLEN SYMMETRISCHEN MATRIX

FROTA ZUR ROTATION DER Q-FAKTOREN

REAL X(NM),Y(NM),Z(NM),V(MM),R(M2MP1),EW(M),
 * CM(MK),CV(MK)
 INTEGER NP(N),KCT(M)
 DIMENSION D(100),CD(100),IND(100)
 IF (M.LE. 100) GOTO 888
 WRITE(6,4004) M
 4004 FORMAT('///' DIE ANZAHL DER VARIABLEN DARF NICHT GROESSER SEIN',
 * ' ALS 100'/' >>QANALY HAT NICHT GERECHNET !!!!!'/')
 RETURN
 888 CONTINUE
 THRESH=0.5
 KE=0
 FM=FLOAT(M)
 DO 10 I=1,M2MP1
 10 R(I)=0.
 SUBTRAHIERE ZEILENWEISE DEN MITTELWERT UND NORMIERE AUF 1.
 DO 50 I=1,N
 SS=0.
 S=0.
 DO 20 L=1,M
 XIL=X(N*(L-1)+I)
 S=S+XIL
 20 SS=SS+XIL**2
 S=S/FM
 SS=SQRT(SS - FM* S**2)
 LK=0
 DO 40 L=1,M
 IL=N*(L-1)+I
 V(L)=(X(IL)-S)/SS
 DO 30 KK=1,L
 LK=LK+1
 30 R(LK)=R(LK)+ V(L)*V(KK)
 40 Z(IL)=V(L)
 50 CONTINUE

BERECNE EIGENWERTE EW(I) UND EIGENVEKTOREN V VON $R=Z^* \times Z$
 CALL EWP (M,R,EW,V)
 CALL EWP (R,V,M,M/2,D,CD,EW,Y,IND)

```

WRITE(6,1000) (EW(L),L=1,M)
DO 70 L=1,M
  IF (EW(L)-THRESH) 80,80,60
60 KE=KE+1
70 CONTINUE
80 WRITE(6,1001) KE
C   BERECHNE HAUPTKOMPONENTEN Y=Z*V VON Z*Z'
DO 100 I=1,N
  DO 100 L=1,KE
    S=0.
    ML=M*(L-1)
    DO 90 J=1,M
      S=S+Z(N*(J-1)+I)*V(ML+J)
      Y(N*(L-1)+I)=S
90
100 CONTINUE
C   MEHRFACHE BERECHNUNG VON ROTIERTEN FAKTOREN (AUFSTIEIGENDE
C   ANZAHL) ZUR BESTIMMUNG MEHRERER (KA-KE) GRUPPIERUNGEN
KA=MINO(K1,K2,KE)
KE=MINO(MAXO(K1,K2),KE)
DO 130 K=KA,KE
  MKK=M*K
C   ROTATION DER Q-FAKTOREN
C   CALL FROTA(N,K,Y)
  CALL VARMX(N,K,Y,NC,R,Z(1),Z(N+1),Z(2*N+1),IER)
  WRITE(6,4711) IER
4711 FORMAT(/'  VARMX - IER =' ,I8)
  WRITE(6,1002)
  DO 110 I=1,N
    WRITE(6,1003) (Y((J-1)*N+I),J=1,K)
110  CONTINUE
C   GRUPPENZUGEHÖRIGKEITEN AUS DEN FAKTORLADUNGEN BESTIMMEN
C   METH=2
METH=3
  CALL ASSIGN(N,K,Y,NP,METH,R,N*K)
C   BERECHNE UND DRUCKE CLUSTERMITTELWERTE
  CALL MEANS(N,M,K,X,NP,CM,CV,KCT,N*M,MKK)
  WRITE(6,1004) K,(KCT(KK),KK=1,K)
  WRITE(6,1005) (NP(I),I=1,N)
  DO 120 L=1,M
120  WRITE(6,1006) (CM(J),CV(J),J=L,MKK,M)
130 CONTINUE
  RETURN
1000 FORMAT(/'  EIGENWERTE: '(1X,8F7.3) )
1001 FORMAT(/'  DARAUS KOENNEN MAXIMAL ',I3,' GRUPPEN GEBILDET WERDEN' /)
1002 FORMAT(/1X,'  Q-FAKTORENLOADUNGEN '/')
1003 FORMAT(1X,10 F 7.3)
1004 FORMAT(/'  CLUSTERZUGEHÖRIGKEITEN'//1X,
* I4,'  CLUSTER DER GROESSEN:',10(2X,I3))
1005 FORMAT(/(1X,15I3) )
1006 FORMAT(/1X,10F8.3)
END

```

```

SUBROUTINE HLCMB1 (N,M,K,X,NP,CM,CV,KCT,SPW,NM,MK)
AUSTAUSCHVERFAHREN (DATENMATRIX IM HAUPTISPEICHER)
PARAMETERLISTE WIE SUBROUTINE MINDS1
REAL X(NM),CM(MK),CV(K)
INTEGER NP(N),KCT(K)
DLARGE=1.E70
IND=0
DO 10 J=1,K
  CV(J)=0.
  KCT(J)=0
  DO 10 L=1,M
    IND=IND+1
    CM(IND)=0.
10  CONTINUE
DO 30 I=1,N
  STANDARDANFANGSPARTITION, FALLS KEINE ANDERE VORGEGEBEN
  IF(NP(I).LT.1.OR.NP(I).GT.K) NP(I)=MOD(I-1,K)+1
  NPI=NP(I)
  KCT(NPI)=KCT(NPI)+1
  V=0.
  IND=(NPI-1)*M
  DO 20 L=1,M
    XH=X(N*(L-1)+I)
    INDL=IND+L
    CM(INDL)=CM(INDL)+XH
    V=V+XH**2
20  CONTINUE
  CV(NPI)=CV(NPI)+V
30  CONTINUE
SPW=0.
DO 50 J=1,K
  DEN=0.
  IF(KCT(J).NE.0) DEN=1./FLOAT(KCT(J))
  V=0.
  IND=(J-1)*M
  DO 40 L=1,M
    INDL=IND+L
    CM(INDL)=CM(INDL)*DEN
    V=V+CM(INDL)**2
40  CONTINUE
  CV(J)=CV(J) - FLOAT(KCT(J))*V
  SPW=SPW+CV(J)
50  CONTINUE
NOSHFT=0
HAUPTSCHLEIFE DER ITERATION
60 CONTINUE
DO 110 I=1,N
  NOSHFT=NOSHFT+1
  IF ( NOSHFT .GT. N ) RETURN
  NPI=NP(I)
  IF ( KCT(NPI) .LE. 1 ) GOTO 110
  V=0.
  IND=(NPI-1)*M
  DO 70 L=1,M
    V=V+(CM(IND+L)-X(N*(L-1)+I))**2
70  CONTINUE
  DEN1=FLOAT(KCT(NPI))
  V1=V*DEN1/(DEN1-1.)
  VMIN=DLARGE
  BESTIMME GRUPPE N=NPI MIT EVTL GRUESSTER VERBESSERUNG
  DO 90 J=1,K

```

```

      IF ( J .EQ. NPI ) GOTO 90
      V=0.
      IND=(J-1)*M
      DO 80 L=1,M
80      V=V+(CM(IND+L)-X(N*(L-1)+I))*2
      DEN2=FLOAT( KCT(J) )
      V2=V*DEN2/(DEN2+1.)
      IF ( V2 .GT. VMIN) GOTO 90
      VMIN=V2
      NPNEU=J
90      CONTINUE
      IF (VMIN .GE. V1 ) GOTO 110
C  GRUPPENWECHSEL UND NEUBERECHNUNG DER GRUPPENCENTROIDE
      NOSHFT=0
      CV(NPI)=CV(NPI)-V1
      CV(NPNEU)=CV(NPNEU)+VMIN
      SPW=SPW-V1+VMIN
      V1=1./(DEN1-1.)
      DEN2= FLOAT( KCT(NPNEU) )
      V2=1./(DEN2+1.)
      IND=(NPI-1)*M
      IND1=(NPNEU-1)*M
      DO 100 L=1,M
      XH=X(N*(L-1)+I)
      INDL=IND+L
      IND1L=IND1+L
      CM(INDL)=(DEN1*CM(INDL)-XH)*V1
      CM(IND1L)=(DEN2*CM(IND1L)+XH)*V2
100      CONTINUE
      NP(I)=NPNEU
      KCT(NPI)=KCT(NPI)-1
      KCT(NPNEU)=KCT(NPNEU)+1
110      CONTINUE
      GOTO 60
      END

```

```

SUBROUTINE KTRANS (N,M,M1,IHK,X,Y,V,R,S,DSIG,EW,TRESH,
* NM,M2MP1,MM)
C   HAUPTKOMONENTENTRANSFORMATION MIT
C   KOVARIANZ- ODER KORRELATIONSMATRIX
C EINGABE:
C N,M   INTEGER   ANZAHL DER ELEMENTE,VARIABLEN
C IHK   REAL      FALLS IHK=0 TRANSFORMATION MIT KOVARIANZMATRIX,
C        REAL      FALLS IHK=1 TRANSFORMATION MIT KORRELATIONSMATRIX
C X     REAL      (N,M)-DATENMATRIX IN 1-DIM SPEICHERUNG
C TRESH REAL      EIGENWERTSCHRANKE FUEER DIE DIMENSION DER TRANSFORMATI
C NM,M2MP1,MM
C M1    INTEGER   DIMENSIONS LAENGEN N*M,M*(M+1)/2,M*M
C AUSGABE:
C M1    INTEGER   ANZAHL DER VARIABLEN IM TRANSFORMIERTEN RAUM
C Y     REAL      (N,M1)-DATENMATRIX NACH TRANSFORMATION
C ARBEITSPARAMETER:
C V     REAL      (M,M)-ARBEITSARRAY IN 1-DIM SPEICHERUNG
C        REAL      IN 1-DIM SPEICHERUNG
C R     REAL      SYMMETRISCHES ARBEITSARRAY IN 1-DIM SPEICHERUNG
C S,DSIG,EW
C        REAL      ARBEITSARRAYS DER LAENGE M
C BENÖTIGTE SUBROUTINE:
C EWP   ZUR BERECHNUNG DER EIGENWERTE UND -VEKTOREN EINER
C        REELLEN SYMMETRISCHEN MATRIX
C
C      REAL V(MM),R(M2MP1),S(M),X(NM),Y(NM),EW(M),DSIG(M)
C      IND=0
C      DO 10 J=1,M
C        S(J)=0.
C        DO 10 L=J,M
C          IND=IND+1
C          R(IND)=0.
C10      CONTINUE
C      DO 30 I=1,N
C        IND=0
C        DO 20 J=1,M
C          XH=X(N*(J-1)+1)
C          S(J)=S(J)+XH
C          DO 20 L=1,J
C            IND=IND+1
C            R(IND)=R(IND)+XH*X(N*(L-1)+1)
C20      CONTINUE
C30      CONTINUE
C      FN=FLOAT(N)
C      DN=1./FN
C      IND=0
C      DO 50 J=1,M
C        S(J)=S(J)*DN
C        IF (IHK .NE. 0 )
C *      DSIG(J)=1./SQRT(R(J*(J+1)/2)*DN-S(J)**2)
C        DO 40 L=1,J
C          IND=IND+1
C          R(IND)=R(IND)*DN-S(J)*S(L)
C          IF ( IHK .NE. 0 ) R(IND)=R(IND)*DSIG(J)*DSIG(L)
C40      CONTINUE
C50      CONTINUE
C LOESUNG DES EIGENWERTPROBLEMS R*V=EW*V
C CALL EWP (M,R,EW,V)
C M1=0
C DO 60 J=1,M
C   IF (EW(J) .LT. TRESH) GOTO 70
C   M1=M1+1
C60  CONTINUE

```

```

70 CONTINUE
  WRITE (6,110) M1,(EW(J),J=1,M)
  DO 80 L=1,M1
    SQEW=1./SQRT(FW(L))
    IND=M*(L-1)
    DO 80 J=1,M
      80   V(IND+J)=V(IND+J)*SQEW
    DO 100 I=1,N
      DO 100 J=1,M1
        SS=0.
        IND=M*(J-1)
        DO 90 L=1,M
          VX=V(IND+L)*(X(N*(L-1)+1)-S(L))
          IF (IHK .NE. 0) VX=VX*DSIG(L)
          SS=SS+VX
        90   CONTINUE
        Y(N*(J-1)+1)=SS
      100   CONTINUE
    RETURN
  110 FORMAT (1X,I2/1X(10F8.2/5X))
  END

```

```

SUBROUTINE MEANS (N,M,K,X,NP,CM,CV,KCT,NM,MK)
  BERECHNUNG DER GRUPPENCENTROIDE UND VARIANZEN AUS X,NP

EINGABE
N      INTEGER      ANZAHL DER ELEMENTE
M      INTEGER      ANZAHL DER VARIABLEN
K      INTEGER      ANZAHL DER CLUSTER
X      REAL         (N,M)-DATENMATRIX IN 1-DIM SPEICHERUNG
NP     INTEGER      FELD DER LAENGE N .ELEMENT I IST IM CLUSTER NP(I)
NM     INTEGER      DIMENSIONS LAENGE N*M
MK     INTEGER      DIMENSIONS LAENGE M*K

AUSGABE
CM     REAL         (M,K)-ARRAY DER MITTELWERTVEKTOREN DER CLUSTER
                        IN 1-DIM SPEICHERUNG
CV     REAL         (M,K)-ARRAY DER VARIANZEN DER VARIABLEN IN
                        DEN EINZELNEN CLUSTERN IN 1-DIM SPEICHERUNG
KCT    INTEGER      FELD DER LAENGE K.KCT(J)=ANZAHL DER ELEMENTE

REAL X(NM),CM(MK),CV(MK)
INTEGER KCT(K),NP(N)
IND=0
DO 10 J=1,K
  KCT(J)=0
  DO 10 L=1,M
    IND=IND+1
    CM(IND)=0.
    CV(IND)=0.
  CONTINUE
10 DO 20 I=1,N
  NPI=NP(I)
  KCT(NPI)=KCT(NPI)+1
  DO 20 L=1,M
    IND1=N*(L-1)+I
    IND=(NPI-1)*M+L
    CM(IND)=CM(IND)+X(IND1)
  CONTINUE
20 DO 40 J=1,K
  IF ( KCT(J) .EQ. 0 ) GOTO 40
  DO 30 L=1,M
    IND=(J-1)*M+L
    CM(IND)=CM(IND) / FLOAT( KCT(J) )
  CONTINUE
30 DO 60 I=1,N
  NPI=NP(I)
  DO 50 L=1,M
    IND=(NPI-1)*M+L
    IND1=N*(L-1)+I
    CV(IND)=CV(IND) +(X(IND1)-CM(IND))**2
  CONTINUE
50 DO 80 J=1,K
  IF ( KCT(J) .EQ. 0 ) GOTO 80
  DO 70 L=1,M
    IND=(J-1)*M+L
    CV(IND)=SQRT( CV(IND)/FLOAT(KCT(J)))
  CONTINUE
70 DO 80 J=1,K
  IF ( KCT(J) .EQ. 0 ) GOTO 80
  DO 70 L=1,M
    IND=(J-1)*M+L
    CV(IND)=SQRT( CV(IND)/FLOAT(KCT(J)))
  CONTINUE
80 RETURN
END

```

```

SUBROUTINE ZTRANS (N,M,X,SM,SS,NM)
C   Z-TRANSFORMATION DER DATENMATRIX X
C   EINGABE:
C   N       INTEGER    ANZAHL DER ELEMENTE (ZEILEN VON X)
C   M       INTEGER    ANZAHL DER VARIABLEN (SPALTEN VON X)
C   X       REAL       (N,M)-DATENMATRIX IN 1-DIM SPEICHERUNG
C                   SPALTENWEISE HINTEREINANDER.
C                   WIRD IM PROGRAMM DURCH Z-TRANSFORMATION
C                   UEBERSCHRIEBEN
C   NM      INTEGER    DIMENSIONSLAENGE VON X: N*M
C   AUSGABE:
C   SM      REAL       MITTELWERTE DER M VARIABLEN
C   SS      REAL       STANDARDABWEICHUNGEN DER M VARIABLEN
C                   DIMENSION X(NM),SM(M),SS(M)
DO 10 J=1,M
    SM(J)=0.
10  SS(J)=0.
DO 30 I=1,N
    DO 20 J=1,M
        IND=(J-1)*N+I
        SM(I)=SM(J)+X(IND)
    20  SS(J)=SS(J)+X(IND)**2
30  CONTINUE
    FN=N
DO 50 J=1,M
    SM(J)=SM(J)/FN
    SS(J)=SQRT( SS(J)/FN - SM(J)**2)
    IND=(J-1)*N
    DO 40 I=1,N
        IDD=IND+I
        X(IDD)=(X(IDD)-SM(J))/SS(J)
    40  CONTINUE
50  CONTINUE
RETURN
END

```

```

C      PROGRAMM KRIT ZUR KRITERIUMSBERECHNUNG
C      KRITERIUM 5.3.3-1
      INTEGER NP1(41),NP2(41)
      REAL K(9)
      N=41
      K1=3
      K2=3
      K1K2=9
      READ(5,1000) (NP1(I),I=1,N)
      WRITE(6,2000) (NP1(I),I=1,N)
10     READ(5,1000,END=99) (NP2(I),I=1,N)
      WRITE(6,2000) (NP1(I),I=1,N)
      WRITE(6,2000) (NP2(I),I=1,N)
      S=SCRIT(N,K1,K2,NP1,NP2,K,K1K2)
      WRITE(6,2003) S
      GOTO 10
99     CONTINUE
      STOP
1000  FORMAT(41I1)
2000  FORMAT(/9X,'NP1=',19I2,5X,10I2,5X,12I2)
2002  FORMAT(/9X,'NP2=',19I2,5X,10I2,5X,12I2)
2003  FORMAT(/9X,'KRITERIUM (5.3.3-1):',F12.3)
      END

```

```

      REAL FUNCTION SCRIT(N,K1,K2,NP1,NP2,K,K1K2)
      INTEGER NP1(N),NP2(N)
      REAL K(K1K2)
      DO 10 LJ=1,K1K2
10     K(LJ)=0.
      DO 20 I=1,N
      LJ= (NP2(I)-1)*K1+NP1(I)
20     K(LJ)=K(LJ)+1.
      S1=0.
      S2=0.
      S3=0.
      DO 40 L=1,K1
      S4=0.
      DO 30 J=1,K2
      S=K((J-1)*K1+L)
      S1=S1+S**2
30     S4=S4+S
40     S2=S2+S4**2
      DO 60 J=1,K2
      IND=(J-1)*K1
      S4=0.
      DO 50 L=1,K1
50     S4=S4+K(IND+L)
60     S3=S3+S4**2
      SCRIT=1.+(S1-(S2+S3)*0.5)/(N*(N-1)*0.5)
      RETURN
      END

```

SPSS-PROGRAMM FÜR DIE SCHRITTWEISE MULTIPLE
DISKRIMINANZANALYSE

```
RUN NAME          AMPLITUDENSPEKTREN
FILE NAME         BIOFEEDBACK
VARIABLE LIST     GRUPPE,FREQ01 TO FREQ40
SUBFILE LIST      A1(41) A2(41) A3(41) A4(41) A5(41) A6(40)
INPUT MEDIUM     OTHER
INPUT FORMAT      FIXED(F4.0,40A4)
RUN SUBFILES      EACH
TASK NAME         DISKRIMINANZANALYSE DER CLUSTER
DISCRIMINANT      GROUPS=GRUPPE(1,3)/
                  ANALYSIS=FREQ01 TO FREQ40/
                  VARIABLES=FREQ01 TO FREQ40/
                  METHOD=WILKS/
                  FUNCTIONS=2,95.,0.05
OPTIONS           3,5,7,10
STATISTICS        1 2 5 6
READ INPUT DATA
FINISH
```

8.1.5 PROGRAMME ZUR VISUALISIERUNG DER LÖSUNGSKURVEN

```

C      PROGRAMM STB  GRAPHISCHE AUSGABE DER LOESUNGSKURVEN
C      PUNKT 6.3.2
C      (ES WERDEN ROUTINEN DES AG-II
C      SYSTEMS (TEKTRONIX) BENUTZT)
C
      INTEGER*2    N1,N2,NARF,IV(60),NPHAS,PHA(60),PHEN(60)
      INTEGER*2 WERTE(12000)
      REAL*4  XH(1000),XL(1000),W1(1000),W2(1000)
      EQUIVALENCE (WERTE(1),W1(1)),(WERTE(4001),W2(1))
      COMMON WERTE
      IIN=5
      IOUT=6
      DO 1 I=1,12000
1  WERTE(I)=0
      INT=300
      READ(5,2000) IANF
      WRITE(IOUT,2001) INT,IANF
      DO 10 IJ=1,400
      CALL DATA (NR,N1,N2,NPHAS,PHA,PHEN,IER,IANF)
      IF (IER.EQ.991) GOTO 999
      WRITE(8) WERTE
      WRITE(IOUT,1001) NR,N1,N2,NPHAS
      WRITE(IOUT,1002) (PHA(I),PHEN(I),I=1,NPHAS)
      NP=NPHAS-1
      DO 10 I=1,NP
      REWIND 8
      READ(8) WERTE
      L=(PHEN(I)*2)+1
      LO=L
      NZ=PHEN(I+1)-PHEN(I)
      IF(NZ.LT.INT) GOTO 25
      DO 30 J=1,NZ
      XL(J)=WERTE(L)
      XH(J)=WERTE(L+1)
      L=L+2
30  CONTINUE
      IPP=I+1
      CALL STABIL (XH,XL,NZ,PHA,PHEN,NR,IPP,W1,W2)
      CALL RTIME(III)
      IF(III.LE.1000) GOTO 999
      GOTO 10
25  CONTINUE
      WRITE(IOUT,1000)NR,IPP,LO,PHEN(IPP),NZ,L
10  CONTINUE
999  WRITE(IOUT,9000)
      STOP
1000 FORMAT(9X,'NR=',I4,'AUSG. IPP=',I3,'< > =',I2I8,'NZ=',I5,
*      ' L(=2*)=',I5)
1001 FORMAT(9X,'NR=',A4,' N1/N2=',I2I8,' NPHAS=',I5)
1002 FORMAT(9X,20I5)
2000 FORMAT(I3)
2001 FORMAT(9X,'STABILITAET ',/9X,'1.PRN-NR (=IANF) =',I5,
1      /9X,'VORGEGEBENES MINIMAL INTERVALL (=INT) =',I5)
9000  FORMAT(9X,'ENDF- MAIN')
      END

```

```

SUBROUTINE DATA (NR,N1,N2,NPHAS,PHA,PHEN,IER,IANF)
C   EINLESEN DER DATEN - BANDVERS.
LOGICAL*1 LW(2,12000)
INTEGER*2 WERTE(12000)
LOGICAL*4 VAPBN
INTEGER*2 N1,N2,NARF,IV(60),NPHAS,PHA(60),PHEN(60),TRJ(60)
INTEGER*2 IREC/0/
EQUIVALENCE (LW(1,1),WERTE(1))
COMMON WERTE
IIN=15
10  READ(IIN,END=20) NR,VAPBN,NARF,IV,NPHAS,
1      (PHA(I),TRJ(I),PHEN(I),I=1,60)
    READ(IIN,END=20) N1,N2,(LW(2,I),I=1,N2)
    IREC=IREC+2
    INR=IREC/2
    WRITE(6,2000) INR,NR,VAPBN,IREC
C   RESTART - FORTLAUFENDE ZAEHLUNG
    IF(INR.LT.IANF) GOTO 10
    IER=0
    RETURN
20  IER=991
    RETURN
2000 FORMAT(9X,'EINGELESEN  INR=',I5,' NR=',A4,' VAPBN ',A4,2X,I5)
END

```

```

SUBROUTINE STABIL (XH,XL,M,PHA,PHEN,NR,IPP,W,W2)
INTEGER*2 PHA(60),PHEN(60)
REAL*4   XH(1000),XL(1000),W(1000),W2(1000)
MA=MINO(1000,M)
C   RR-INTERVALL => FREQ.
DO 10 I=1,MA
10  W(I)=100./XH(I)
    W2(1)=0.0
    DO 20 I=2,MA
    W2(I)=W2(I-1)+XH(I)
20  CONTINUE
C   AUFRUF DER AG-II ROUTINEN
    CALL INITT
    CALL CHR3:Z(3)
    CALL ERASE
    CALL BELL
    CALL TXAGII (IPP,MA,NR,PHEN,PHA)
C   UEBERGEHEN DER PHASE =G(GO)=71
    CALL TINPUT(IC)
    IF(IC.EQ.1) GOTO 33
    CALL BILD12 (W,W2,XL,MA)
    CALL ERASE
    CALL TXAGII (IPP,MA,NR,PHEN,PHA)
    MA1=MA-1
    DO 11 I=1,MA1
11  XL(I)=XL(I+1)
    CALL BILD3 (W,XL,MA1)
33  CONTINUE
C   BLANK=32
    DO 22 J=1,10
    CALL BELL
    CALL TINPUT(IC)

```

```

      IF ( IC.NE. 32 ) RETURN
C 22 CONTINUE
      E=69 (E:DBEDINGUNG)
      IF(IC.NE.69) RETURN
      CALL ADUTST(8,'ENDE TCS')
      CALL BELL
      WRITE(6,516)
516 FORMAT(9X,'ENDE TCS')
      CALL FINITT
      STOP
      END

```

```

SUBROUTINE TXAGII (IPP,MA,NR,PHEN,PHA)
  INTEGER*2 PHA(60),PHEN(60)
  REAL*8 TEXT(4),XNR,XPH
  IPHE=PHEN(IPP-1)
  IPH2=PHEN(IPP)
  IPHAS=PHA(IPP)
  CALL BCNV(IPP,TEXT(1),103,8,0)
  CALL BCNV(IPHE,TEXT(2),103,8,0)
  CALL BCNV(IPH2,TEXT(3),103,8,0)
  CALL BCNV(MA,TEXT(4),103,8,0)
  CALL BCNV(NR,XNR,103,8,0)
  CALL BCNV(IPHAS,XPH,103,8,0)
  CALL HOME
  CALL ADUTST(10,'          NR=')
  CALL ADUTST(8,XNR)
  CALL ADUTST(8,'          PHA= ')
  CALL ADUTST(8,XPH)
  CALL ADUTST(8,'          IPP*=')
  CALL ADUTST(8,TEXT(1))
  CALL ADUTST(8,' PHEN< >')
  CALL ADUTST(8,TEXT(2))
  CALL ADUTST(8,TEXT(3))
  CALL ADUTST(8,'          MA=')
  CALL ADUTST(8,TEXT(4))
  RETURN
  END

```

```

SUBROUTINE BILD12 (W,W2,XL,MA)
  REAL*4 W(1000),W2(1000),XL(1000)
  CALL BINITT
  CALL NPPTS(MA)
  CALL SLIMX(500,4000)
  CALL SLIMY(2000,3000)
  CALL CHECK(W2,W)
  CALL DISPLAY(W2,W)
  CALL BINITT
  CALL NPPTS(MA)
  CALL SLIMX(500,4000)
  CALL SLIMY(500,1500)
  CALL LINE(4)
  CALL CHECK(W2,XL)
  CALL DISPLAY(W2,XL)
C  HARD COPY =H=72
  CALL BELL
  CALL TIK-UT (IC)

```

```

IF (IC.EQ.72) CALL HDCOPY
RETURN
END

```

C

```

SUBROUTINE BILD3 (X,Y,N)
REAL*4 X(1000),Y(1000)
CALL BINITT
CALL SLIMX(500,4000)
CALL SLIMY(200,3000)
CALL STEPS(3)
CALL SYMRL (-10)
CALL NPTS(N)
CALL CHECK(X,Y)
CALL DSPLAY(X,Y)
CALL RSET US
HARDCOPY =H=72
CALL BELL
CALL TINPUT (IC)
IF (IC.EQ.72) CALL HDCOPY
RETURN
END

```

4711

```

SUBROUTINE USERS (X,Y,IS)
LOGICAL FIRST/.TRUE./
REAL CD/0.0707/
REAL XA,YA
REAL SALPHA,CALPHA
CD=1.
WRITE(6,*) FIRST,X,Y,IS
IF (FIRST) GOTO 4711
CALPHA=SQRT((X-XA)**2+(Y-YA)**2)
WRITE(6,*) CALPHA
IF (CALPHA.EQ.0.) RETURN
SALPHA=(Y-YA)/CALPHA
CALPHA=(X-XA)/CALPHA
WRITE(6,*) SALPHA,CALPHA
CALL DRAWR(-CD*(CALPHA-SALPHA),-CD*(CALPHA+SALPHA))
CALL MOVER(-2*CD*SALPHA,2*CD*CALPHA)
CALL DRAWR(CD*(CALPHA+SALPHA),-CD*(CALPHA-SALPHA))
CONTINUE
XA=X
YA=Y
FIRST=.FALSE.
RETURN
ENTRY RSET US
FIRST=.TRUE.
RETURN
END

```

8.2 ANMERKUNGEN

ANMERKUNGEN ZU PUNKT 1

- 1/1 Die Psychophysiologie wird hier also als ein relativ neues und interdisziplinäres Fachgebiet verstanden, in dem abgrenzbar zur weitergefaßten physiologischen Psychologie (vgl. u.a. DEUTSCH & DEUTSCH 1973; GROSSMAN 1973 oder THOMPSON 1975) der Versuch einer Integration von psychologischen Inhalten bzw. Elementen mit Elementen aus der Physiologie, spezieller der Neurologie, sodann mit Elementen aus der (biologischen) Kybernetik und Informationstheorie unternommen wird und in dem der Einsatz modernerer Mittel der Elektronik bzw. Halbleitertechnik im Unterschied zu anderen Sektoren der Psychologie dominiert. Würde man wissenschaftshistorisch umstandslos auf die Forschungsergebnisse des 19. Jahrhunderts etwa von BELL, MÜLLER, DU BOISREYMOND, YOUNG, HELMHOLTZ, HERING oder WEBER rekurren (siehe die Aufnahme und Weiterentwicklung in WUNDT 1874) und diese als entscheidende und allein geschichtlich relevante Bezugsgrößen für die Psychophysiologie angeben, dürfte man die von sehr divergierenden Aspekten aus zu diskutierende Problematik nur unzureichend reflektieren, welche eine an einem systemtheoretisch-funktionalen Paradigma (KUHN 1967, bes. 28ff.) orientierte Wissenschaft beinhaltet, deren charakterische Züge auch auf der Folie einer tendenziellen "Finalisierung" (BÖHME et al 1972; 1973, insbes. 129) gesehen und metawissenschaftlich analysiert werden können (vgl. in diesem Kontext zur Resistenz und Rezeptivität VAN DEN DAELE & WEINGART 1975).
- 1/2 KEUPP (1972a und b) und die dort genannten weiteren Arbeiten, in denen unter Aufnahme historischer und sozialer Implikationen des medizinischen Krankheitsmodells mehr theoretisch akzentuierte "Umriss für eine sozialwissenschaftliche Alternative zum 'Krankheitsmodell'" (SCHEFF in KEUPP 1972, 136ff.) zu finden sind, welche allerdings in der angedeuteten dichotomisierenden Auffassung über sozial- versus naturwissenschaftliche Zugriffsweisen und Methoden nur partiell die eigentliche Problematik der Anamnese, Diagnose und Therapie komplex verankerter Störungen in Angriff nehmen und einer Lösung zu führen.
- 1/3 Vgl. in Bezug auf PICKENHAIN (1968) und LANG (1973) BIRBAUMER (1973, 2 sowie 1974, 226f.); hier die Erweiterung von LEGEWIE & NUSSELT (1975, 6f.).
- 1/4 Indikatoren sind insbesondere die Arbeiten zur Aktivierungsforschung (DUFF 1962; HAIDER 1967 oder SCHÖNPFUG 1969; genereller GREENFELD & STERNBACH 1972 oder VENABLES & CHRISTIE 1975).

- 1/5 Daß weitergehende Differenzierungen in einzelne Relevanzarten wie etwa "axiomatische, technologische oder soziale Relevanz" (SPIEGEL - RÖSING 1973, 86ff.) fragwürdig sind und die Kategorie der Relevanz (vgl. HOLZKAMP 1972, 288ff.) prinzipieller hinterfragbar ist, dürfte augenfällig sein. Dennoch wird hier dieser Terminus verwandt, um einer möglichen Reduktion von Wissenschaftstheorie auf Methodologie vorzubeugen, die zu einer sehr unterschiedlichen Beurteilung der Relevanz "therapeutischer Interventionen" beitragen könnte.
- 1/6 Evident scheint bereits die Flut einschlägiger Artikel zu sein. Vgl. als Übersichtsversuch BROWN (1975), ebenfalls die Sammelbände von BARBER et al (1970-1974), neuerdings auch HUME (1976), im deutschen Sprachraum z.B. LEGEWIE & NUSSELT (1975) und BIRBAUMER (1978).
- 1/7 Gleichsam wegweisend waren die in den sechziger Jahren an der Rockefeller Universität von dem Team um N.E. MILLER geleisteten Studien. Obwohl deren Replizierbarkeit nicht gelang (vgl. die Replik von MILLER & DWORKIN 1974), dürfte ihnen dennoch die nicht unwichtige Funktion zuzuschreiben sein, die These von der uneingeschränkten "Autonomie des vegetativen Systems" für einen breiteren Leserkreis problematisiert zu haben. Neuere Arbeiten weisen darauf hin, daß in der Regel nicht einzelne, isolierte physiologische Funktionen, sondern eher bestimmte Funktionszusammenhänge oder physiologische "Muster" über Prozesse somatisch autonomer Integration operant konditionierbar sind (siehe u.a. OBRIST 1976).
- 1/8 Da eine unmittelbare corticale Regulation "autonomer" Funktionen sehr unwahrscheinlich ist und ausgeschlossen werden kann, kommt den afferenten und efferenten Bahnen vom Cortex zu hypothalamischen und limbischen Strukturen besondere Bedeutung zu. So werden corticofugale Impulse, welche über den Hypothalamus und das limbische System verschaltet sind, für die operante Konditionierbarkeit vegetativer Funktionen verantwortlich gemacht. Experimentell wird diese Annahme durch Decortisierungsversuche unterstützt, in denen auch nach massiver intrakranieller positiver Verstärkung keine diesbezüglichen Lernerfolge zu registrieren waren (DICARA et al 1970; grundlegend HAYMAKER et al 1969 oder HOCKMANN 1972; als verständliche Einführung in das vegetative Nervensystem JÄNIG 1976).
- 1/9 Als Übersicht über die Biofeedback-Forschung unter speziellem Bezug auf die Stabilisation der Herzfrequenz VAITL (1975a); generell neuerdings BIRBAUMER (1978).
- 1/10 Siehe hierzu die vom Bundesminister für Jugend, Familie und Gesundheit editierten Materialbände zur psychiatrischen Versorgung in der BRD (bes. Bd. 10).

- 1/11 VAITL (1975a, 192).
- 1/12 Abgesehen von Problemen der Messung psychophysiologischer Größen (vgl. BIRBAUMER 1974, 233f.) besteht die Gefahr artifizieller Ergebnisse nicht zuletzt durch eine unreflektierte Anwendung uni- und bivariater Verfahren der (linearen) Statistik. Arbeiten, in denen entweder mathematisch-statistische Aspekte in Hinsicht auf die Analyse psychophysiologischer Signale aufgegriffen oder Modelle der Frequenzanalyse mit Verfahren der multivariaten Statistik methodisch-systematisch aufeinander bezogen werden, finden sich bislang in der Psychophysiologie selten, wohingegen Fragen der Meßtechnik und -aufnahme vergleichsweise intensiv behandelt werden. Als Indikator für die Tendenz, Stellenwert und Funktion mathematischer Modelle für Prozesse der Konstruktion psychophysiologischer Theorien und Hypothesen zu vernachlässigen, mag auch die kürzlich publizierte Darstellung von LANC (1977, 32ff.) über "Speicherung und Verarbeitung (Hervorh.d.Verf.) psychophysiologischer Signale" gewertet werden (vgl. speziell die Verweise auf S. 39). Über den allerdings mehr generellen Einsatz des Computers in der Klinischen Psychologie z.B. HARRISON (1973), ROBINSON (1973) oder BUCHSBAUM & COPPOLA (1974).
- 1/13 Als Orientierung siehe die Übersichten in LEGEWIE & NUSSELT (1975, 74f.) oder ECKERT et al (1978, 1434ff.).
- 1/14 Zur Psychophysiologie der Angst etwa SPIELBERGER (1972), BIRBAUMER (1973), ZUCKERMAN & SPIELBERGER (1976) oder LEVITT (1976); zum Streß SELYE (1956; 1960), BROADBENT (1971), SARASON & SPIELBERGER (1975); zur Motivation EWERT (1965), LEGEWIE & EHLERS (1972), GROSSMAN (1973) oder THOMPSON (1975). Vgl. in diesem Kontext auch die Hypothese von der "Kontrolle autonomer Funktionen durch Konditionierung konkordanter Muster", wie sie etwa von BIRBAUMER (1978, 2090f.) referiert wird, und die sich ergebenden methodischen Konsequenzen.
- 1/15 Als eine Art gegenläufige Tendenz, die hier nur anzudeutende und beim Einsatz multivariater Modelle auftauchende Problematik der Iso- bzw. Homomorphie numerischer und empirischer Relative lediglich am Rande zu erwähnen, kann die in jüngerer Zeit einsetzende Diskussion über die Angemessenheit formaler Modelle gewertet werden, welche zunehmend die Funktion der Kritik anstelle der Kodifizierung übernimmt und in der Psychologie sich um das faktorenanalytische Modell konzentriert. Vgl. bereits KALLINA (1967), KALVERAM (1970a-d, bes. 1970d, 229), auch KEMPF (1972), auf das Dilemma der (mathematischen) Psychologie verweisend KEMPF (1974, 18), als Beispiele auch BUSE & PAWLIK (1978) oder GIGERENZER & STRUBE (1978).

- 1/16 Vgl. hierzu BIRBAUMERS "methodische Vorbemerkungen" (1973, 2) sowie die bereits genannte Arbeit von VAITL (1975a).
- 1/17 So wird das EEG in Biofeedback-Versuchen etwa bei der Therapie von Schlafstörungen und Formen der Epilepsie eingesetzt (siehe die Zusammenstellung in BROWN 1975, 409ff.; deutsch LEGEWIE & NUSSELT 1975, 103ff. und BIRBAUMER 1978, 2086ff.).
- 1/18 So zu Beginn der EEG-Forschung und klinischen EEG-Diagnostik Ende der zwanziger Jahre H. BERGER.
- 1/19 Speziell NEUNDÖRFER (1975, 17f.); allgemeiner zum EEG GIBBS & GIBBS (1971) oder COOPER et al (1974, bes. 161ff.) einführend SCHMIDT (1976, 152ff.); aus psychologischer Sicht PAWLIK (1974).
- 1/20 Vgl. zu diesen Ausführungen die Arbeit von KÜNKEL (1972), hier S. 18 .
- 1/21 Zur Zeitbereichsdarstellung des EEGs HJORTH (1970; 1973), ebenso DENOTH (1975).
- 1/22 Neben KÜNKEL (1972) siehe aus der Vielzahl der Arbeiten in neuerer Zeit VOS (1975), DRESCHER & FERBER (1976), FERBER & EICHHOLZ (1976) oder CHAVANCE & SMASON-DOLLFUS (1976).
- 1/23 Zu diesem Punkt vgl. genauer die oben angegebene Literatur. In konventionellen EEG-Prozessoren (etwa dem BIO 16 von AEG-TELEFUNKEN) werden diese Parameter automatisch berechnet und ausgegeben.
- 1/24 KÜNKEL (1972) bzw. konkreter die Punkte 3 und 4 dieser Arbeit. Abb. 1.3-1 ist ebenfalls an eine Darstellung in KÜNKEL (1972, 19) angelehnt.
- 1/25 Die notwendigen Verkürzungen, insbesondere der mathematischen Grundlagen, sollen zumindest partiell durch Literaturverweise aufgehoben werden. Siehe im übrigen auch den publizierten Anhang "Grundbegriffe aus der Mengenlehre und Linearen Algebra" in STEINHAUSEN & LANGER (1977, 185ff.).

ANMERKUNGEN ZU PUNKT 2

- 2/1 VAITL (1975a, 192ff.).
- 2/2 Siehe die Ausführungen in dem Punkt 1.3 dieser Arbeit.
- 2/3 Es wird also weder Vollständigkeit noch Repräsentativität angestrebt, zumal dieses angesichts der sich progredient verkürzenden Verdoppelungszeiten wissenschaftlicher Produktion auch ein fundamentales Mißverständnis andeuten würde (vgl. näher z.B. DE SOLLA PRICE 1974, bes. 2off.). Das Ziel dieser Arbeit besteht vielmehr darin - wie einleitend betont -, sowohl bekannte als auch bislang nur sehr selten oder gar nicht benutzte Ansätze und Methoden zur psychophysiologischen Signalanalyse nach einheitlichen Kriterien zu systematisieren und für einen breiteren Leserkreis einführend zu beschreiben.
- 2/4 Da auf Meßmethoden in der Psychophysiologie üblicherweise häufig und ausführlich eingegangen wird, sei an dieser Stelle auf eine entsprechende Darstellung verzichtet (vgl. im deutschen Sprachraum LANG 1973, LANG 1977 oder für die Klinische Psychologie insgesamt ECKERT et al 1978).
- 2/5 Siehe die Übersichten in LEGEWIE & NUSSELT (1975, 74f.) und ECKERT et al (1978, 1434ff.) neben den einschlägigen Büchern der Physiologie.
- 2/6 Die Aufzeichnung wurde dankenswerterweise von Herrn Dr.med. R. KORINTHENBERG, Universitätsklinik Münster, zur Verfügung gestellt. Eine äquivalente bzw. vergleichbare Simultanregistrierung physiologischer Variablen stellt im Rahmen der Biofeedback-Forschung das sog. "Summenbiofeedback" (BIRBAUMER 1978, 2094ff.) dar.
- 2/7 Vgl. einführend HEINECKER 1975.
- 2/8 Neben bekannten angloamerikanischen Monographien und Lehrbüchern hierzu im deutschen Sprachraum etwa UNBEHAUEN (1969) oder UNBEHAUEN & BAUER (1974), neuerdings die ausgezeichnete Arbeit von WINKLER (1977).
- 2/9 Unterstellt wird also, daß keine deterministischen Beziehungen vorliegen. Über den benutzten Begriff der Zufallsvariable oder zufälligen Veränderlichen vgl. die einführenden Kapitel in den Lehrbüchern zur Wahrscheinlichkeitstheorie und Statistik, z.B. FISZ (1965, 27ff.), HENGST (1967, 31ff.) oder MESCHKOWSKI (1968, 106ff.).
- 2/10 Siehe genauer die Darstellungen in den erwähnten Lehrbüchern, speziell FISZ (1965, 228ff.); aus mathematischer Sicht auch in neuerer Zeit LAMPARTI (1977) und LIPTSER & SHIRYAYEV (1978). Selbstverständlich handelt es sich hier um vereinfachte Unterscheidungen; so ist speziell (2.2-6) eine mathematische Idealisierung, da unendlich lange Signale in der Psychophysiologie streng genommen nicht existieren.

- 2/11 Vgl. FISZ (1965, 269ff.) sowie die anschließende Definition eines stationären Prozesses im weiteren Sinne.
- 2/12 Vgl. FISZ (1965, 217) bzw. die unter Anm. 2/9 zitierten Arbeiten.
- 2/13 Zur Identifizierung der Periodenlänge siehe auch den interessanten Ansatz von LAMPRECHT & WEBER (1972).
- 2/14 Während psychologische Phänomene und Prozesse nur sehr bedingt zu formalisieren sind, gelingt dies aufgrund der Gegenstandsdifferenz in der Physiologie bzw. Medizin allgemein in zunehmenden Maße. Charakteristische Approximationsaufgaben, die zumeist mittels exponentieller Ausdrücke bearbeitet werden, stellen sich in der Pharmakokinetik, wenn etwa die In- und Exvasion eines (oral) applizierten Pharmakons untersucht wird (BATEMAN-Funktionen), oder beispielsweise in der Strahlentherapie im Rahmen der Entwicklung optimaler Therapiepläne (exponentielle vers. sigmoide Dosiswirkungskurven etc.). Einen eher forschungsrelevanten Ansatz, dem ebenfalls ein exponentielles Modell zugrundeliegt, stellen HOKE & HIEKE (1976) auf dem Gebiet der Elektrochleographie vor.
- 2/15 Gemeint ist der Approximationssatz von WEIERSTRASS (WERNER & SCHABACK 1972, 87ff.). Neben diesem Lehrbuch liegen den folgenden Ausführungen die Ausarbeitungen von COLLAT & KRABS (1973) sowie WERNER (1966) zugrunde.
- 2/16 Vgl. im folgenden auch KANTOROWITSCH & AKILOW (1978) und weiterführend TRIEBEL (1972).
- 2/17 Über Vollständigkeit und Separabilität (Cauchysche Konvergenzkriterium) einführend KANTOROWITSCH & AKILOW (1978, Ioff.).
- 2/18 Über den vereinfachend an diesem Punkt vorausgesetzten Existenzsatz vgl. WERNER & SCHABACK (1972, 86ff.).
- 2/19 Aufgrund der Monotonie führt m.a.W. (2.5.2.2-8) zum gleichen Resultat wie (2.5.2.2-9).
- 2/20 Siehe wiederum KANTOROWITSCH & AKILOW (1978, 71f.) oder TRIEBEL (1972, 85).
- 2/21 Darüberhinaus besitzen sie die Eigenschaft, sogenannte Eigenfunktionen linearer (zeitinvarianter) Systeme zu sein (vgl. WINKLER 1977 oder ACKERMANN 1972). Die numerische Berechnung dieser Funktionen wurde desweiteren durch die Entwicklung effektiver Algorithmen erheblich vereinfacht, so daß sie auch unter ökonomischen Gesichtspunkten in der Praxis vorteilhaft sind (vgl. Punkt 3.2.3).

- 2/22 WALSH (1923). Als verständliche und auf praktische Fragen zielende Einführung ist BEAUCHKAMP (1975) zu nennen. Die zahlreichen Anwendungen und Entwicklungstendenzen sind u.a. den jährlichen Proceeding-Bänden "Applications of Walsh Functions" zu entnehmen. Die Definition (2.5.3.2-1) wird von HARMUTH (1972) vorgeschlagen.
- 2/23 Im folgenden BEAUCHKAMP (1975, 12ff.).
- 2/24 Diese Beziehung ist von praktischem Interesse, da man auf einer digitalen Anlage die Fourierkoeffizienten auch unter Ausnutzung dieser Relation bedeutend schneller als auf dem direkten Weg berechnen kann.
- 2/25 COURANT & HILBERT (1968, 74f.).
- 2/26 a.a.O. .
- 2/27 Diese Transformation wird in der Literatur auch einfach als z-Transformation im Unterschied zur FISHERschen Z-Transformation bezeichnet.
- 2/28 Vgl. einführend COURANT & HILBERT (1968, 26ff.) .
- 2/29 Als Standardartikel WATANABE (1967) und die zahlreichen Arbeiten in den IEEE Trans.Comp.; ein Anwendungsbeispiel auf dem medizinischen Sektor etwa wird von BRUGGER (1973) gegeben.
- 2/30 Zu Spline-Funktionen siehe das Lehrbuch von BÖHMER (1974) oder die Darstellung in WERNER & SCHABACK (1972, 154ff.), aus pragmatischer Sicht SPÄTH (1973); über neuere Entwicklungen neben dem von BÖHMER et al (1976) editierten Kongreßbericht WERNER (1977a, b).
- 2/31 Hierzu die eben genannten Arbeiten, speziell BÖHMER (1974, 176ff.).
- 2/32 Vgl. unter Angabe von Beispielen SPÄTH (1973).
- 2/33 Bei wachsendem r gleicht sich der Spline immer mehr der gegebenen Funktion an, bis er für $r=m$ (d.h. die Knotenzahl ist gleich der Zahl der Punkte der Funktion) zum interpolierenden Spline wird.
- 2/34 Nach REINSCH (1967).
- 2/35 ALGOL-Programme werden in BÖHMER (1974, 282ff.) und FORTRAN-Routinen in SPÄTH (1973) bzw. in entsprechenden Programmbibliotheken angegeben (siehe empfehlenswert die Subroutinen in der NAG-Library).

ANMERKUNGEN ZU PUNKT 3

- 3/1 Auf diesen in der Praxis eher vernachlässigten Punkt, auf den schon WIENER (1949) hingewiesen hatte, wird zumindest indirekt durch die rasche Verbreitung der Walsh-Funktionen und deren spektrale Interpretation aufmerksam gemacht (siehe (2.5.3.2) bzw. einführend BEAUCH-KAMP 1975, bes. 85ff.).
- 3/2 Dazu aus zum Teil rein mathematischer Sicht PAPOULIS (1962), BRACEWELL (1965) oder deutsch LIGHTHILL (1966); anwendungsorientiert zur spektralen Analyse bzw. zur sog. Zeitreihenanalyse neben dem Standardwerk von BLACKMAN & TUKEY (1959) JENKINS & WATTS (1968), BOX & JENKINS (1970) oder ANDERSON (1971), empfehlenswert OPPENHEIM & SCHAFER (1975) oder RABINER & GOLD (1975); einführend in die Fourieranalyse von Zeitreihen BLOOMFIELD (1976) und CHATFIELD (1975); über die schnelle Fouriertransformation (FFT) instruktiv speziell für Sozialwissenschaftler BRIGHAM (1974).
- 3/3 Die Ausführungen können auch für allgemeinere Voraussetzungen gelten, worauf hier jedoch nicht eingegangen werden soll.
Für die folgende Formel (3.2.1-2) ist vereinfachend angenommen, daß die Funktion $s(t)$ zu jedem Zeitpunkt t stetig ist.
- 3/4 Siehe zu diesem Punkt ausführlicher die einschlägigen Kapitel über Integralgleichungen und -transformationen in den Lehrbücher der Höheren Mathematik.
Integraltransformationen sind wichtige Lösungsmethoden etwa für gewöhnliche und partielle Differential-, Integral- oder Differenzgleichungen. Man nutzt hier das aus der schulischen Logarithmenrechnung bekannte Prinzip, daß die gesuchte Lösung einfacher dadurch zu erzielen ist, indem man die Aufgabe transformiert, dort eine entsprechende ("äquivalente"), in der Regel kennzeichnenderweise wesentlich einfachere Operation ausführt und durch Rücktransformation die gewünschte Lösung erhält. Neben der Fouriertransformation und der Walsh-Transformation spielt in der Praxis vor allem die Laplace-Transformation eine wichtige Rolle, welche beispielsweise in der Elektrotechnik zur Prüfung linearer Schaltkreise berechnet wird. Zur folgenden Darstellung siehe SMIRNOV (1964, 422ff.).
- 3/5 Die obige Bedingung (3.2.1-1) schränkt - obgleich sie nicht notwendig ist - die Anwendung der Fouriertransformation in der Praxis ein, indem etwa Potenz-, Exponential- oder konstante Funktionen nicht berücksichtigt werden. In naheliegender Weise läßt sich durch Erweiterung des Konvergenzbegriffs im Sinne der distributiven Konvergenz die

verallgemeinerte Fouriertransformierte einführen für beliebige Funktionen (Distributionen). Über den weiteren, nicht uninteressanten Zusammenhang der Fouriertransformation zu anderen Integraltransformationen, speziell der Laplace-Transformation, siehe die entsprechenden Lehrbücher bzw. Bücher aus der mathematischen Physik.

- 3/6 Weitere Eigenschaften (Differentiation, Fouriertransformierte einer Ableitung u.s.w.) sind der Literatur zu entnehmen bzw. können abgeleitet werden. In dieser Arbeit geht es vorwiegend um eine einführende, auf sozialwissenschaftliche Anwendungen zielende Darstellung.
- 3/7 Vgl. COOLEY et al (1969) und den historischen Exkurs in BRIGHAM (1974, 8f.).
- 3/8 Anstelle der Basis 2 wird bei diesen Algorithmen mit der Basis 4, 8, 16 oder einer gemischten Basis gearbeitet (BERGLAND 1968; 1969a) siehe auch SINGLETON 1969).
- 3/9 Ein effektiver Algorithmus für $m \neq 2^r$ wird von SINGLETON (1969) vorgestellt. In der Praxis ist darauf zu achten, daß die Primfaktoren aus Gründen des Rechenaufwands klein sind (< 5). In SINGLETON (1969, 99) findet sich eine entsprechende Tabelle.
- 3/10 Verständlich und unter Angabe von FORTRAN-Programmen wiederum BRIGHAM (1974, bes. 148ff.). Vergleichsweise gut ausgetestete und zuverlässige Subroutinen zur FFT enthält das Programmsystem NAG (Subroutinen Co6).
- 3/11 Ausführlicher die unter 3/2 auszugsweise zitierte Literatur. Im Unterschied zu dem Symbol für die Faltungsoperation meint nachfolgend die Bezeichnung $S(f,P)$ die konjugiert komplexe Form.
- 3/12 Gemeint ist die FFT (3.2.3). Nach Berechnung des Leistungsspektrum durch Fouriertransformation erhält man m.a.W. durch einfache Bildung der inversen Fouriertransformation die gesuchte Autokorrelationsfunktion. Hardwaremäßig wird damit der ansonsten erforderliche Korrelator überflüssig. Analog wird man die Berechnung der Kreuzkorrelation vereinfachen. Auf diese Aspekte wird explizit in Punkt 3.6 eingegangen. Zum Wiener-Khinchine Theorem siehe auch die Darstellung in WINKLER (1977, 179ff.).
- 3/13 Wie an den anderen Textstellen auch sind an die Funktionen z und h geeignete Integrierbarkeitsvoraussetzungen zu stellen, also z.B. z, h stetig.

- 3/14 Die zugehörige Impulsantwort lautet dann

$$h(t) = (2f_0 \sin(2\pi f_0 t)) / (2\pi f_0 t) \quad .$$

 Anstelle des idealisierten Tiefpaßfilters, für den die Bedingung (3.5-2) entfällt ($h(t) \neq 0$ für $t < 0$) würde man in der Praxis zur on-line Analyse einen entsprechenden Filter durch Festsetzung geeigneter Frequenzbänder und der Berechnung der zugehörigen Übertragungsfunktionen realisieren.
- 3/15 Vgl. LIDTHILL (1966, 19ff.).
- 3/16 In einer solchen Situation würde man durch Summation der Einzelsignale die gewichtete Mittelung erhalten. Obiger Filter ist im übrigen auf nicht periodische Signale anwendbar, indem man diese durch ein synchron gesetztes zweites Signal (fast) periodisch abtastet (vgl. einführend SPRENG & KEIDEL 1971; SPRENG 1972 bzw. DAVID et al 1971 und die zahlreichen Arbeiten speziell aus der Audiometrie).
- 3/17 Auf die verschiedenen einzelnen Filtertypen (Inverser-, Vorhersage- oder etwa Optimal-Filter) wird nicht eingegangen. Siehe aus der Vielzahl von Arbeiten die Darstellung in WINKLER (1977, 430ff.), speziell die KALMAN-BUCY Filterung, die auch gesonderter Gegenstand der Arbeit von BRAMMER & SIFFLING (1975) ist und speziell im medizinischen Sektor zunehmend Verwendung findet.
- 3/18 Anstelle der bisherigen Integral- würde man also an dieser Stelle die jeweilige Summenschreibweise benutzen und die diskrete Form numerisch auswerten.
- 3/19 Die Herleitung des Abtasttheorems aus dem Faltungssatz (3.3-2) entnehme man der Literatur (vgl. bes. die instruktive Darstellung in WINKLER 1977, 67ff.). Dieser Zusammenhang ist im übrigen etwa seit den frühen zwanziger Jahren aus der Theorie der Integralgleichungen bekannt und erlangte erst 1949 durch die Wiederaufnahme in den SHANNON'schen Informationstheorie seine eigentliche Bedeutung.
- 3/20 Als verständliche Einführung etwa JENKINS & WATTS (1968, 239ff.), grundlegend WELCH (1967) und von Interesse weiterhin EBERHARD (1973).
- 3/21 Vgl. als ausgesprochen informative Übersicht BERGLAND (1969b), insbesondere die Abbildung 3 (S. 43) neben den unter (3.2.3) genannten Arbeiten.

ANMERKUNGEN ZU PUNKT 4

- 4/1 VAITL (1975a, 192ff.)
- 4/2 Inwieweit ein intraindividueller Vergleich zur Beantwortung der obigen Zielfragen, insbesondere der Effekt-Generalisation zuverlässige und valide Kriterien liefert, dürfte ein bislang allenfalls im Ansatz gelöstes Problem darstellen (zur Einzelfalldiagnostik im deutschen Sprachraum HUBER 1973 und 1978, bes. den dortigen historischen Exkurs S. 1154ff.; vgl. auch GOTTMANN 1973; in Bezug auf die Biofeedback-Forschung BARLOW et al 1977). Einzelfallstudien dürfte eher eine "hypothesengenerierende Funktion" (VAITL 1975b, 205) zukommen. Siehe aus wissenschaftstheoretischer Perspektive generell zum Validierungskonzept u.a. FELTER (1977).
- 4/3 Als geradezu "klassisches" dimensionsstatistisches Verfahren in der Psychologie darf die Faktorenanalyse gelten (HORST 1965; HARMAN 1967; LAWELY & MAXWELL 1971; deutsch PAWLIK 1968; ÜBERLA 1971; WEBER 1974 oder REVENS DORF 1977). Neuere faktorenanalytische Modelle etwa von TUCKER, McDONALD oder JÖRESKOG sind jedoch bisher in der Forschungspraxis nur vereinzelt zu finden (zum Konzept der "konfirmatorischen Faktorenanalyse" WEEDE & JAGODZINSKI 1977). Im Rahmen der Mustererkennung werden entsprechende Ansätze auch allgemeiner unter den Begriff "Feature Selection" subsumiert (einführend MEISEL 1972, 162ff.; FU 1976, von Interesse insbes. GRENANDER 1977 und KLINGER et al 1977). Auf den Zusammenhang von Faktorenanalyse und problemabhängiger Reihenentwicklung bzw. Parameterextraktion ist in (2.5.4) hingewiesen worden.
- 4/4 Zur Differenzierung von Diskrimination und Klassifikation bzw. Gruppierung auch KENDALL (1966, 165), KENDALL & STUART (1968, 314), SKARABIS (1970, 2) oder BOCK (1974, 23), in Bezug auf die Klinische Psychologie BAUMANN (1974). Grundlegend zur automatischen Klassifikation und damit Clusteranalyse das wohl als Standardwerk zu bezeichnende Buch von BOCK (1974), praxisnah ANDERBERG (1973) etc.; ohne daß jeweils explizit die Beziehung deutlich herausgestellt werden kann zu den folgenden Ausführungen LANGER (1975a, b) sowie STEINHAUSEN & LANGER (1977).
- 4/5 Diese Formel wird speziell für die hierarchischen Verfahren der Clusteranalyse noch von praktischer Bedeutung sein (Punkt 4.3.2).
- 4/6 Einführend WEINBERG (1973), informativ KOONTZ et al (1975) und der dort angegebene Algorithmus sowie die zahlreichen Artikel in den IEEE Trans.Comp. zu diesem Thema bzw. ebenso zur Dynamischen Programmierung.

- 4/7 Die mit der Skalierung verbundenen Probleme werden an dieser Stelle als bekannt vorausgesetzt, so daß auf eine an sich erforderliche Diskussion der Entsprechung empirischer und numerischer Strukturen (Iso- bzw. Homomorphieproblem) verzichtet wird (siehe auch heute noch von gewisser Aktualität SUPPES & ZINNES 1963; als elementare Einführung in diese Problematik KRIZ 1973, bes. 26ff. und die dortigen pointierten Formulierungen).
- 4/8 Vgl. hierzu die Definitionen in (2.5.2.1).
- 4/9 Ausführlicher BOCK (1974, 77ff.).
- 4/11 Formel (4.2.2-3) wird auch "Canberra-Metrik" und (4.2.2-4) "nicht-metrischer" Koeffizient genannt. Letzterer hat jedoch den Nachteil, nicht skaleninvariant zu sein (LANCE & WILLIAMS 1967b).
- 4/12 Siehe genauer SKARABIS (1970) und BOCK (1974, 63ff.) über probabilistische Ähnlichkeitsmaße; vgl. in diesem Zusammenhang ebenso die informationstheoretischen Definitionsansätze (BOCK 1974, 93ff.).
- 4/13 Als informative und übersichtliche Zusammenstellung wiederum BOCK (1974, 66ff.).
- 4/14 Diese Annahme ist nicht unproblematisch, da bei vielen korrelierenden Variablen die Matrix K oft nicht invertierbar ist. In diesem Fall würde man bei der Bildung von U und $\text{diag}(\lambda_i^{-1/2})$ nur nicht verschwindende Eigenwerte und zugehörige Eigenvektoren berücksichtigen.
- 4/15 Vgl. ausführlicher BOCK (1974, 40ff.).
- 4/16 Diese Metrik ist insbesondere bei Lage- und Standortproblemen von praktischer Relevanz, etwa bei der Ermittlung der optimalen Position elektronischer Bauelemente aber auch von Kindergärten, Bushaltestellen oder Notrufsäulen in rechtwinkligen Straßennetzen. Ein einfacher Algorithmus zur L_1 -Clustering wird von SPÄTH (1976) wiedergegeben.
- 4/17 Vgl. BOCK a.a.O. .
- 4/18 Nicht selten werden clusteranalytische Verfahren im Sinne eines "Zwei-Phasen" Modells aufgefaßt (JARDINE & SIBSON 1971, xi; RAO 1972, 17; RAND 1971, 846 oder LANCE & WILLIAMS 1966, 373).
- 4/19 An dieser Stelle sei bereits auf die interessante methodische Alternative der nicht-exhaustiven Konfigurationsfrequenzanalyse (KFA) hingewiesen (Punkt 4.3.4.2).

- 4/20 Im folgenden auch BOCK(1974, 383ff.).
- 4/21 Nach LANCE & WILLIAMS (1966, 374). Die Tabelle ist entnommen aus STEINHAUSEN & LANGER (1977, 77). Siehe ebenso die weitere Diskussion der Eigenschaften der verschiedenen Strategien, insbesondere der graphentheoretischen Version des "Single-Linkage" Algorithmuses (S. 94ff.).
- 4/22 EDWARDS & CAVALLI-SFORZA (1965) bzw. zum Algorithmus GOWER (1967)
- 4/23 Auf Methoden und Programme zur Erzeugung geeigneter Startpartitionen und den damit verknüpften Aspekten kann nur verwiesen werden (siehe u.a. STEINHAUSEN & LANGER 1977, 137).
- 4/24 Vgl. die unter 4/44 zitierten Arbeiten, speziell ANDERBERG (1973) aus pragmatischer Sicht.
- 4/25 In der multivariaten Statistik wird diese Testgröße üblicherweise als WILKsches Lambda bezeichnet (COOLEY & LOHNES 1971,
- 4/26 Näher zu diesen Zusammenhängen BOCK (1974, bes. 184ff.).
- 4/27 Dieses Kriterium wurde u.a. von HOTELLING (1954) als Testgröße verwandt und dürfte ebenso wie das WILKsche Lambda als bekannt vorausgesetzt werden, so daß auf weitere Erläuterungen verzichtet werden kann.
- 4/28 Für den Fall $r=2$ etwa stellt der Gruppencentroid und für $r=1$ der Vektor der Mediane unter präzisierten Bedingungen ein entsprechendes Klassenzentrum dar.
- 4/29 Dieses an sich naheliegende Verfahren wurde auch von FORGY (1965) vorgeschlagen. Vgl. die Ausführungen in BOCK (1974, 171ff.).
- 4/30 Siehe STEINHAUSEN & LANGER (1977, 107).
- 4/31 RUBIN (1967). Treffender wäre allerdings die Bezeichnung 'hill-descending'.
- 4/32 Programme und Beispiele hierzu finden sich in zahlreichen Artikeln und Büchern. Als neueren Beitrag, allerdings ausgesprochen heuristisch und zum Teil verkürzend SPÄTH (1977a).
- 4/33 Vgl. STEINHAUSEN & LANGER (1977, 135ff.).
- 4/34 In der Regel wird an dieser Stelle auf STEPHENSON rekurriert, der in der Mitte der 30iger Jahre in Anlehnung

an THOMSON von der 'Q-Technik' spricht (auch 'inverted' oder 'transposed' factor technique; vgl. OBERLA 1971, 297f.). Angemessener wäre anstelle der Verwendung des Begriffs 'Technik' die Charakterisierung der jeweils zugrundegelegten Datenebene, da eine Differenzierung etwa in S, P, T oder Q-Technik bei Invarianz des involvierten Modells lediglich eine Etikettierung beinhaltet (siehe PAWLIK 1968, 279ff.).

- 4/35 Siehe die Darstellung z.B. von JESDINSKY (1971).
- 4/36 Die einzelnen Publikationen finden sich zusammengefaßt in KRAUT & LIENERT (1973). Eine elementare Einführung versucht ROEDER (1974).
- 4/37 Unter Verzicht auf Information können unschwer ordinal- oder intervall-skalierte Daten auf mehrkategoriale nominale Daten abgebildet werden. Praktische Probleme tauchen in der Regel dadurch auf, daß die Anzahl der pro Merkmalzelle vorhandenen Fälle meist zu gering ist (s.u.). Dies läßt sich vermeiden, indem man bestimmte Kategorien nach präzisierten und inhaltlich schlüssigen Kriterien zusammenfaßt.
- 4/39 Ausführlicher zur KFA KRAUTH & LIENERT (1973); zu Programmen STEINHAUSEN & LANGER (1977, 152ff.); über die Typisierung von Verlaufskurven mit Hilfe der KFA-Prinzipien BARTOSZYK & LIENERT (1978).
- 4/40 Hierzu wiederum BOCK (1974, 420f.) und das Programm von FABER & NOLLAUS (1969), das entgegen des relativ eleganten Ansatzes nur für sehr kleine Fallzahlen mit einem vertretbaren Rechenaufwand zu realisieren ist und von dort eher für wissenschaftsinterne Zwecke zu verwenden sein dürfte.
- 4/41 Vgl. die einführende Darstellung in (2.5.5.2); als interessante Arbeit BONEVA et al (1971).
- 4/42 BOCK (1974, 249ff.).
- 4/43 Siehe die Arbeiten auf den Gebieten der Mustererkennung, Künstlichen Intelligenz oder Dynamischen Programmierung, speziell der Branch-and-Bound Verfahren, sowie die Publikationen in IEEE Trans.Comp., Comp.Journal, Comm.ACM, Pattern Recognition, Comp.Bio.med.Res., Biometrics bzw. anderen Periodika sowie die einschlägigen Kongreßberichte.
- 4/44 JARDINE & SIBSON (1971), ANDERBERG (1973), BOCK (1974), als allerdings verkürzter Überblick DURAN & ODELL (1974), aus sozialwissenschaftlicher Sicht VOGEL (1975), SCHLOSSER (1976), einführend SODEUR (1974), als Algorithmensammlung HARTIGAN (1975) neben ANDERBERG (1973), pragmatisch ausgerichtet SPÄTH (1975) oder STEINHAUSEN & LANGER (1977) und die dort zitierten weiteren Arbeiten, als Fallstudiensammlung SPÄTH (1977b).

ANMERKUNGEN ZU PUNKT 5

- 5/1 So die Einschätzung von BIRBAUMER (1975, 165). Zu problematischen Aspekten auf diesem Gebiet VAITL (1975a).
- 5/2 Vgl. neben VAITL & KENKMANN (1972), VAITL (1975a) neuerdings VAITL et al (1978).
- 5/3 Siehe VAITL et al (1978, 1).
- 5/4 a.a.O. S. 7; allgemeiner auch BIRBAUMER (1978, 2084) bzw. zur Frage eines (durch Konditionierung konkordanten) Musters S. 2090f. .
- 5/5 Vgl. ausführlicher VAITL et al (1978).
- 5/6 Zur Methodik und zum Versuchsaufbau siehe genauer VAITL & KENKMANN (1972, bes. 254ff.).
- 5/7 Vgl. die weiteren Erläuterungen und Abbildungen a.a.O. .
- 5/8 VAITL et al (1978, Fig.2).
- 5/9 Zur Problematik des Biofeedback am Beispiel der Herzfrequenzkontrolle VAITL (1975a) und die dort zitierten Arbeiten; als Übersichtsversuch BROWN (1975, 201ff.) bzw. stark verkürzt neuerdings auch BIRBAUMER (1978, 2088ff.); zur Therapie speziell VAITL (1975b) und die dort zu findende Relativierung hinsichtlich der therapeutischen Anwendbarkeit (S. 217); zu eher theoretischen Aspekten OBRIST (1976); als einzelne Studien in jüngerer Zeit etwa HARRISON & RASKIN (1976), BOUCHARD & CORSON (1976), LANG & TWENTYMAN (1976), LACEY & LACEY (1977), BLANCHARD et al (1977); zu dem nicht uninteressanten Diskriminationsphänomen CLEMENS & Mac DONALD (1975), McFARLAND (1975) oder EPSTEIN et al (1977). Nicht allein die Zahl der auf diesem Sektor publizierten Arbeiten ist als Indikator anzusehen, daß die "Formulierung präziser Gesetzmäßigkeiten" (BIRBAUMER 1978, 2084) ebenso hier nur sehr eingeschränkt möglich ist.
- 5/10 Vgl. in Bezug auf die RR-Distanzen LUCZAK & LAURIG (1973), SAYERS (1973) oder MULDER & MULDER-HAJONIDES VON DER MEULEN (1973); zu "Basis-Parametern" etwa ROHMERT et al (1973, 37); über Verteilungen JENNINGS et al (1974); zur Kombination von Frequenz- und Faktorenanalyse OPMEER (1973) Aus dieser fragmentarischen Übersicht läßt sich bereits äußerlich die Bedeutung der Untersuchung der Herzrate für ergonomische Fragestellungen erkennen (vgl. u.a. BOYCE 1974).

- 5/11 Vgl. u.a. LUCZAK & RASCHKE (1975).
- 5/12 Im folgenden LUCZAK & RASCHKE (1975, 2), siehe auch LUCZAK & LAURIG (1973).
- 5/13 Vgl. zur Blutdruckregulation etwa HILDEBRANDT (1967) bzw. einschlägige Lehrbücher der Physiologie.
- 5/14 Vorausgesetzt wird, daß die Atemtätigkeit sich in physiologischen Grenzen sich befindet. Es wird ein ungefährender Mittelwert von 14 Zügen pro Minute angenommen (THEWES 1976, 486), der in Phasen der Entspannung allerdings geringer sein dürfte. Siehe zum Phänomen der respiratorischen Arrhythmie die tierexperimentelle Arbeit von KOEPCHEN & THURAU (1959) oder ANGELONE & COULTER (1964), in Bezug auf die Herztätigkeit ebenso ENGEL et al (1969), speziell SROUFE (1971), im Rahmen der Diskrimination HART (1975) und die dort angegebene Korrektur; über dynamische Wechselbeziehungen zwischen Atmung und cardiovascularem System den Entwurf von WIRTH (1973).
- 5/15 Vgl. VAITL et al (1978, 4ff.).
- 5/16 Diese Einschränkung ist aus der Thematik dieser Arbeit zwingend abzuleiten und didaktisch legitimiert und wird der an sich möglichen direkten Berechnung der Diskriminanzanalyse von dort vorgezogen, die man im Rahmen des Gesamtprojekts an diesem Punkt zutreffenderweise berechnen würde.
- 5/17 Gemeint ist die nicht in allen Fällen glückliche, d.h. angemessene Unterscheidung in einem Merkmals- und einen Klassifikationsraum.
- 5/18 Vgl. STEINHAUSEN & LANGER (1977, 169ff.) bzw. die unter Anmerkung 4/44 aufgeführten Arbeiten.
- 5/19 Diese auch als C^* -Kriterium bezeichnete Funktion wurde u.a. von RAND (1971) vorgeschlagen, besitzt allerdings den Nachteil, daß keine Signifikanzaussagen bzw. abschätzungen verfügbar sind.
- 5/20 Numerisch zweckmäßiger und auch in der Subroutine SCRIT implementiert ist die folgende Umformung des Ausdrucks (5.3.3-1)

$$s(G, G') = 1 - \frac{1}{n(n-1)} \left(\sum_{i=1}^k \left(\sum_{j=1}^1 n_{ij} \right)^2 + \sum_{j=1}^1 \left(\sum_{i=1}^k n_{ij} \right)^2 - 2 \sum_{i=1}^k \sum_{j=1}^1 n_{ij}^2 \right).$$

Im übrigen scheitert der an sich plausible unmittelbare Vergleich der Zielfunktionswerte (wird ein Verfahren aus 4.3.3 vorausgesetzt) daran, daß Lösungen mit gleichen Zielfunktionswerten nicht notwendig ähnliche oder gar gleiche Klasseneinteilungen besitzen.

- 5/21 Vgl. die bereits mehrfach genannte Arbeit von VAITL et al (1978) als weiteren Bezugsrahmen.
- 5/22 Die Berechnungen wurden auf der nahezu total ausgelasteten Maschine des Rechenzentrums der Universität Münster (IBM 360/50) durchgeführt. Aufbauende und interessante zusätzliche numerische Vergleiche konnten aus diesem Grund nicht realisiert werden.
- 5/23 Insgesamt wurden 311 Versuchspersonen in der Phase der Datenorganisation bearbeitet, von denen jeweils bestimmte Gruppen speziellen experimentellen Bedingungen unterworfen waren.
- 5/24 Vgl. HULL & DOBELL (1962) oder DIETER & AHRENS (1971), instruktiv KNUTH (1969, 10ff.; bes. 103f.).
- 5/25 Die Zahl der Probanden reduzierte sich hierdurch auf insgesamt 296. Von den für diese Arbeit zur Verfügung stehenden Personen mußte ein Proband ausgeschlossen werden (Gruppe B)
- 5/26 Auf genauere Angaben wird verzichtet, da die Recordstruktur dem Programmtext zu entnehmen ist (Punkt 8.2.1).
- 5/27 Durch diese Programmstruktur wurde im übrigen die Kompatibilität mit dem vom Verfasser entwickelten Clusteranalyse-Programmen erreicht (vgl. 8.2.4).
- 5/28 Siehe u.a. CHATFIELD (1975, 148). Dies kann im übrigen auch aus (3.2.3) gefolgert werden.
- 5/29 Die Koeffizienten werden rekursiv mittels der Subroutine FKOEFF berechnet (SSP-Manual Vers.III, 275).
- 5/30 Siehe WOMACK (1971, 400) oder den Vergleich in LUCZAK & LAURIG (1973).
- 5/31 Vgl. hierzu auch die Arbeit von SPÄTH (1977a).
- 5/32 Die Abkürzungen entsprechen den benutzten Routinen, die unter (8.2.4) aufgelistet sind.
- 5/33 Zur Frage der Startpartition und (zufälligen)"Standardanfangspartition" vgl. STEINHAUSEN & LANGER (1977, bes. 125f).
- 5/34 Siehe die Ausführungen in Punkt 2.5.4 .

- 5/35 Aus didaktischen Gründen wird m.a.W. auf die Auflistung der übrigen Ergebnisse verzichtet, die im übrigen anhand der benutzten Programme unschwer selbst nachvollzogen werden können. Die Diskriminanzanalyse wurde nach dem im Vergleich zu anderen Programm sehr effektiven und zugleich transparenten SPSS-Programm DISCRIMINANT berechnet. Die zugehörigen Parameterkarten finden sich in (8.2.4).
- 5/36 Da es sich in dieser Arbeit um eine an einem Beispielsatz aufzuzeigende Integration von Methoden der Frequenzbereichs darstellung und multivariaten Modellen unter methodischer Schwerpunktsetzung handelt, wird der an sich indizierte intraindividuelle Vergleich nicht gezogen.
- 5/37 Man wird in einer anderen Situation die Einzelergebnisse genauer analysieren und etwa graphisch darstellen und die Cluster bezüglich passiver, d.h. nicht zur Klassifikation benutzter physiologischer und psychologischer Variablen auswerten.

ANMERKUNGEN ZU PUNKT 6

- 6/1 Aus der Vielzahl einschlägiger Lehrbücher COLLATZ (1967) oder KAMKE (1977).
- 6/2 Im folgenden KAMKE (1977), dem auch die betreffenden Eindeutigkeits- und Existenzsätze sowie deren Beweise zu entnehmen sind, die hier vereinfachend vorausgesetzt werden.
- 6/3 Vgl. a.a.O. , Teil A .
- 6/4 a.a.O., insbes. S. 47ff. .
- 6/5 Siehe hierzu die Ausführungen in (2.5.4) oder (4.3.4.1).
- 6/6 KAMKE (1977, bes. 42ff.) bzw. in Bezug auf systemtheoretische Betrachtungen ACKERMANN (1972, 193ff.) oder HARTMANN & DELF (1974, 56ff.), instruktiv SCHWARZ (1976).
- 6/7 Über die Differenzierung zwischen der 1.Methode und der 2. oder direkten Methode von Ljapunov siehe ACKERMANN (1972), HATRMANN & DELF (1974) und SCHWARZ (1976); neben LA SALLE & LEFSCHETZ (1967) neuerdings in diesem Zusammenhang ROUCHE et al (1977); im folgenden SCHWARZ (1976, 44ff.).
- 6/8 Oft wird die euklidische Norm (2.5.2.1-7) in der Praxis zugrundegelegt.
- 6/9 Vgl. wiederum SCHWARZ (1976, 44ff.).
- 6/10 Siehe COLLATZ (1967, 42ff.), zur Tabelle S. 45; einführend ebenso KAMKE (1977, 42ff.).
- 6/11 Gemeint sind Ansätze der Art von LUCZAK & RASCHKE (1975).
- 6/12 Sicherlich stellt ein DGL-System, wie es in (6.2.2.3) angegeben ist, eine hinsichtlich der Herz- und Atemfunktionen und deren Interdependenzen eine sehr grobe Vereinfachung dar, zumal ersichtlich ist, daß aus der in (3.5-2) beschriebenen Realisierbarkeitsbedingung folgend ein Korrekturglied (etwa in Form eines speziellen Integralterms) in die Systemgleichung einbezogen werden sollte. Insbesondere für Arbeiten auf dem psychophysiologischen Sektor, welche den analytischen Vorteil systemtheoretischer Konstruktion und Identifikation zu nutzen suchen, dürften regelungstheoretische Ansätze zukünftig größere Chancen besitzen, wie sie in Bezug auf ergonometrische Fragestellungen etwa von LUCZAK & RASCHKE (1975) publiziert worden sind. Interessante Analysen dieser Form finden sich u.a. in den Bilog.Cybernetics, Math.Biosciences und speziell in den Lecture Notes in Biomathematics.

9 LITERATURVERZEICHNIS

- ACKERMANN, J. Abtastregelung. Berlin 1972
- AHRENS, H. J. Multidimensionale Skalierung. Weinheim 1972
- ANDERBERG, M. R. Cluster Analysis for Applications. New York 1973
- ANDERSON, T. W. An Introduction to Multivariate Statistical Analysis. New York 1958
- The Statistical Analysis of Time Series. New York 1971
- ANGELONE, A. & COULTER, N. A. Respiratory Sinus Arrhythmia: A Frequency Dependent Phenomenon. J. Appl. Physiol. 19, 1964, 541-557
- BARBER, T. X. et al (Eds.) Biofeedback and Selfregulation. Chicago 1970/74
- BARLOW, D. H. et al Single-Case Design and Clinical Biofeedback Experimentation. Biofeedback and Selfregulation 2, 1977, 221-239
- BARTOSZYK, G. D. & LIENERT, G. A. Konfigurationsanalytische Typisierung von Verlaufskurven. Z. exp. angew. Psych. 25, 1978, 1-9
- BAUMANN, U. Gruppierung und Klassifikation: statistische Probleme in der Klinischen Psychologie. In: SCHRAML, W. & BAUMANN, U. (Eds.) Klinische Psychologie II. Bern 1974, 77-112
- BEAUCHKAMP, K. G. Walsh Functions and their Applications. London 1975
- BECKMANN, D. et al Methodenprobleme in der Psychotherapieforschung. In: PONGRATZ, L. J. (Ed.) 1978, 1085-1118
- BERGER, H. Über das Elektroencephalogramm des Menschen. Arch. Psychiat. Nerven. 87, 1929, 527-570
- BERGLAND, G. D. A Fast Fourier Transform Using Base 8 Iterations. Math. Comput. 22, 1968, 275-279
- A Radix-Eight Fast Fourier Transform Subroutine for Real-Valued Series. IEEE Trans. Audio. Electroacoust. AU-17, 1969, 138-140 (1969a)

- A Guides Tour of the Fast Fourier Transform.
IEEE Spectrum 6, 1969, 41-52 (1969b)
- BIRBAUMER, N. (Ed.) Neuropsychologie der Angst. München 1973
- BIRBAUMER, N. Zur Anwendung psychophysiologischer Methoden
in der Verhaltensmodifikation. In: KRAIKER, C.
(Ed.) Handbuch der Verhaltenstherapie. Mün-
chen 1974, 225-254
- Physiologische Psychologie. Berlin 1975
- Biofeedback. In: PONGRATZ, L.J. (Ed.) 1978,
2083-2103
- BLACKMANN, R.B. & TUKEY, J.W. The Measurement of Power Spectra. New York
1959
- BLANCHARD, J.M. et al The Use of Feedback-Training and a Stimulus
Control Procedure to Obtain Large Magnitude
Increases in Heart Rate outside the Labora-
tory. Biofeedback and Selfregulation 2, 1977,
81-91
- BLOOMFIELD, P. Fourier Analysis of Time Series: An Intro-
duction. New York 1976
- BOCK, H.H. Automatische Klassifikation. Göttingen 1974
- BÖHME, G. et al Alternativen in der Wissenschaft. Z.f.Soz.
1, 1972, 302-316
- Die Finalisierung der Wissenschaft. Z.f.Soz.
2, 1973, 128-144
- BÖHMER, K. Spline Funktionen. Stuttgart 1974
- BÖHMER, K. et al (Eds.) Spline Functions. Berlin 1976
- BONEVA, L.I. et al Spline Transformation - Three New Diagnostic
Aids for the Statistical Analyst. J. Royal
Stat. Society, Series B, 33, 1971, 1-37
- BOUCHARD, C. & CORSON, J.A. Heart Rate Regulation With Success and Fail-
ure Signals. Psychophysiology 13, 1976, 69-74
- BOX, G.E.P. & JENKINS, G.M. Time Series Analysis: Forecasting and Control.
San Francisco 1970
- BOYCE, P.R. Sinus Arrhythmia as a Measure of Mental Load.
Ergonomics 17, 1974, 177-183
- BRACEWELL, R. The Fourier Transform and its Applications.
New York 1965

- BRAMMER, K. & SIFFLING, G. Kalman-Bucy-Filter. München 1975
- BRIGHAM, O.E. The Fast Fourier Transform. New Jersey 1974
- BROADBENT, D.E. Decision and Stress. New York 1971
- BROWN, B. The Biofeedback Syllabus - A Handbook for the Psychophysiologic Studie of Biofeedback. Springfield 1975
- BROWN, P.B. Data Analysis Techniques. In: BROWN, P.B. et al (Eds.) Electronics for Neurobiologists. Cambridge 1973, 418-481
- BRUGGER, P. Die Anwendung der Karhunen-Loeve Entwicklung im Hinblick auf die Klassifizierung von Elektroencephalogrammen. Dipl.Arbeit TU-München 1973
- BUCHSBAUM, M. & COPPOLA, R. Computer Use in Bioelectric Data Collection and Analysis. In: THOMPSON, R.F. & PATTERSON, M.M. (Eds.) Bioelectric Recording Techniques. Part B. New York 1974, 210-244
- BUSE, L. & PAWLIK, K. Stichprobenfehler und Rotationsambiguität von Faktorenladungen. Z.exp.angew.Psych. 25, 1978, 10-34
- CHATFIELD, C. The Analysis of Time Series: Theory and Practice. London 1975
- CHAVANCE, M. & SAMSON-DÖLLFUS, D. A Contribution by Multidimensional Methods to Choosing Description Data for EEG by Spectral Analysis. Meth.Inform.Med. 15, 1976, 229-236
- CHIEN, Y.T. & FU, K.S. On the Generalised Karhunen-Loeve Expansion. IEEE Trans.Inform.Theory 11, 1967, 518-520
- CLEMENS, W.J. & Mac DONALD, D.F. Relationship between Heart Beat Discrimination and Heart Rate Control. Psychophysiology 13, 1976, 176
- COLLATZ, L. Differentialgleichungen. Stuttgart 1967³
- COLLATZ, L. & KRABS, W. Approximationstheorie. Stuttgart 1973
- COOLEY, J.W. et al Historical Notes on the Fast Fourier Transform. IEEE Trans.Audio.Electroacus. AU-15, 1969, 76-79
- COLLEY, J.W. & TUKEY, J.W. An Algorithm for the Machine Calculation of Complex Fourier Series. Math.Comput. 19, 1965, 297-301

- COOLEY, W.W. & LOHNES, P.R. Multivariate Data Analysis. New York 1971
- COOPER, R. et al Elektroenzephalographie. Frankfurt 1974
- COURANT, R. & HILBERT, D. Methoden der mathematischen Physik. Berlin 1968³
- DAVID, E. et al Korrelationsanalyse von gemittelten und ungewichteten evozierten Potentialen. Kybernetik 9, 1971, 22-26
- DENOTH, F. Some General Remarks on Hjorth's Parameters Used in EEG Analysis. In: DOLCE, G. & KÜNKEL, H. (Eds.) Computerized EEG Analysis. Stuttgart 1975, 9-18
- DE SOLLA PRICE, D.J. Little Science, Big Science. Frankfurt 1974
- DEUTSCH, J.A. & DEUTSCH, D. Physiological Psychology. Homewood 1973
- DICARA, L.V. et al Classical Conditioning and Instrumental Learning of Cardiac and Gastrointestinal Responses Following Removal of Neocortex in the Rat. J. comp.physiol.Psych. 73, 1970, 208-216
- DIETER, U. & AHRENS, J. An Exact Determination of Serial Correlations of Pseudo-Random Numbers. Numer.Math. 17, 1971, 101-123
- DRESCHER, D. & FERBER, G. Dreidimensionale Darstellung von segmentweise berechneten EEG-Spektren. Meth.Inform.Med. 15, 1976, 90-93
- DUFFY, E. Activation and Behavior. New York 1962
- DURAN, B.S. & ODELL, P.L. Cluster Analysis. A Survey. Berlin 1974
- EBERHARD, A. An Optimal Discrete Window for the Calculation of Power Spectra. IEEE Trans.Audio.Electroacous. AU-21, 1973, 37-43
- ECKERT, A.F. et al Grundlagen elektronischer Meßtechnik in der Klinischen Psychologie. In: PONGRATZ, L.J. (Ed.) 1978, 1428-1470
- EDWARDS, A.W. & CAVALLI-SFORZA, L.L. A Method for Cluster Analysis. Biometrics 21, 1965, 362-375
- ENGEL, P. et al Der Tagesgang der Phasenkoppelung zwischen Herzschlag und Atmung in Ruhe und seine Beeinflussung durch dosierte Arbeitsbelastung. Int. Z. angew. Physiol. 27, 1969, 339-355

- EPSTEIN, L.H. et al Heart Rate as a Discriminative Stimulus. Psychophysiology 14, 1977, 143-149
- EWERT, O. Gefühle und Stimmungen. In: THOMAE, E. (Ed.) Handbuch der Psychologie, 2. Band: Motivation. Göttingen 1965, 229-271
- FABER, E. & NOLLAU, W. Über ein Verfahren zur automatischen Klassifikation. DRZ Heft S-6. Darmstadt 1969
- FELTER, F. Zur Validierung theoretischer Konstrukte - Wissenschaftstheoretische Aspekte des Validierungskonzepts. Z.f.Soz. 6, 1977, 370-385
- FERBER, G. & EICHHOLZ, G. Distribution of Variables Extracted from Spectral EEG-Analysis. Meth.Inform.Med. 15, 1976, 94-98
- FISZ, M. Wahrscheinlichkeitsrechnung und Mathematische Statistik. Berlin (VEB) 1965^o
- FORGY, E.W. Cluster Analysis of Multivariate Data: Efficiency versus Interpretability of Classifications (abstract.). Biometrics 21, 1965, 768
- FU, K.S. (Ed.) Digital Pattern Recognition. Berlin 1976
- GIBBS, F.A. & GIBBS, E.L. Elektroencephalographie. Stuttgart 1971
- GIGERENZER, G. & STRUBE, G. Zur Revision der üblichen dimensionsanalytischer Verfahren. Z.Entw.psych.Päd.Psych. 10, 1978, 75-86
- GOTTMAN, J.M. N-of-One and N-of-Two Research in Psychotherapy Psychol.Bull. 80, 1973, 93-105
- GOWER, J.C. A Comparison of Some Methods of Cluster Analysis. Biometrics 23, 1967, 623-638
- GRENNANDER, U. Pattern Analysis. Berlin 1978
- GREENFELD, N.S. & STERNBACH, R.A. (Eds.) Handbook of Psychophysiology. New York 1972
- GROSSMAN, S.P. Essentials of Physiological Psychology. New York 1973
- HAIDER, M. Elektrophysiologische Indikatoren der Aktiviertheit. In: SCHÖNPFUG, W. (Ed.) 1969, 125-156
- HARMAN, H.H. Modern Factor Analysis. Chicago 1967²

- HUBER, H.P. Psychometrische Einzelfalldiagnostik. Weinheim 1973
- Kontrollierte Fallstudie. In: PONGRATZ, L.J. (Ed.) 1978, 1153-1199
- HULL, T.E. & DOBELL, A.R. Random Number Generator. Siam Rev. 4, 1962, 230-254
- HUME, W.I. Biofeedback Therapy. Montreal 1976
- JÄNIG, W. Das vegetative Nervensystem. In: SCHMIDT, R.F. & THEWS, G. (Eds.) 1976, 114-145
- JARDINE, N. & SIBSON, R. Mathematical Taxonomy. New York 1971
- JESDINSKY, H. Mehrdimensionale Kontingenztafeln. In: WALTER, E. (Hg.) Statistische Methoden II. Berlin 1970
- JENKINS, G.M. & WATTS, D.G. Spectral Analysis and its Applications. San Francisco 1968
- JENNINGS, J.R. et al A Comparison of the Statistical Distribution of Beat-by-Beat Heart Rate and Heart Period. Psychophysiology 11, 1974, 207-210
- KALLINA, H. Das Unbehagen in der Faktorenanalyse. Psychol. Beiträge 10, 1967, 81-86
- KALVERAM, K.T. Über Faktorenanalyse. Arch.Psychol. 122, 1970, 92-118 (1970a)
- Probleme der Selektion in der Faktorenanalyse I. Arch.Psychol. 122, 1970, 199-214 (1970b)
- Probleme der Selektion in der Faktorenanalyse II. Arch.Psychol. 122, 1970, 215-222 (1970c)
- Probleme der Selektion in der Faktorenanalyse III. Arch.Psychol. 122, 1970, 223-230 (1970d)
- KAMKE, E. Differentialgleichungen. Lösungsmethoden und Lösungen. Bd. I. Stuttgart 1977⁹
- KANTOROWITSCH, L.W. & AKILOW, G.P. Funktionalanalysis in normierten Räumen. Frankfurt 1978
- KEMPF, W.F. Zur Bewertung der Faktorenanalyse als psychologische Methode. Psychol.Beiträge 14, 1972, 610-625
- Dynamische Modelle zur Messung sozialer Verhaltensdispositionen. In: KEMPF, W.F. (Ed.) Probabilistische Modelle in der Sozialpsychologie. Bern 1974, 13-55

- KENDALL, M.G. Discrimination and Classification. In: KRISHNAIAH, P.R. (Ed.) Multivariate Analysis. New York 1966, 165-185
- Time Series. London 1973
- KENDALL, M.G. & STUART, D. The Advanced Theory of Statistics. Vol. III. London 1968²
- KEUPP, H. (Ed.) Der Krankheitsmythos in der Psychopathologie. München 1972 (1972a)
- Psychische Störungen als abweichendes Verhalten. München 1972 (1972b)
- KLINGER, A. et al (Eds) Data Structures, Computer Graphics, and Pattern Recognition. New York 1977
- KNUTH, D.E. The Art of Computer Programming. Vol.2: Semi-numerical Algorithms. Massachusetts 1969
- KOEPCHEN, H.P. & THURAU, K. Über die Entstehungsbedingungen der atem-synchronen Schwankungen des Vagustonus (Respiratorische Arrhythmie). Pflug.Arch.Physiol. 269, 1959, 10-30
- KOONTZ, W.L.G. et al A Graph-Theoretic Approach to Nonparametric Cluster Analysis. IEEE Trans.Comp.C-25, 1976, 936-944
- KOWALSKI, H.J. Lineare Algebra. Berlin 1975⁷
- KRAUTH, J. & LIENERT, G.A. KFA-Konfigurationsfrequenzanalyse. Meisenheim 1973
- KRIZ, J. Statistik in den Sozialwissenschaften. Reinbeck 1973
- KÜNKEL, H. Die Spektraldarstellung des EEG. Z.EEG-EMG 3, 1972, 15-24
- KUHFITTIG, P.K.F. Introduction to the Laplace Transform. New York 1978
- KUHN, T.S. Die Struktur wissenschaftlicher Revolutionen. Frankfurt 1967
- LACEY, B.C. & LACEY, J.I. Change in Heart Rate Period: A Function of Sensomotor Event Timing within the Cardiac Cycle. Physiol.Psychol. 5, 1977, 383-393
- LAMPARTI, J. Stochastic Process - A Survey of the Mathematical Theory. New York 1977
- LAMPRECHT, G. & WEBER, F. Eine neue Methode zur Bestimmung von Periodenlängen. Pflug.Arch.Physiol. 336, 1972, 60-72

- LANC, O. Psychophysiologische Methoden. Stuttgart 1977
- LANCE, G.H. & WILLIAMS, W.T. A General Theory of Classificatory Sorting Strategies I. Hierarchical Systems. Comp.J.9, 1966, 373-380
- A General Theory of Classificatory Sorting Strategies II. Clustering Systems. Comp.J.10, 1967, 271-277 (1967a)
- Mixed Data Classificatory Programm I. Austr. Comp.J. 1, 1967, 15-20 (1967b)
- LANDGRAF, C. Stochastische Lineare Systeme. TU-Berlin, Reihe Brennpunkt Kybernetik, Berlin 1976
- LANG, P.J. Die Anwendung psychophysiologischer Methoden in der Psychotherapie und Verhaltensmodifikation. In: BIRBAUMER, N. (Ed.) 1973, 11-79
- LANG, P.J. & TWENTYMAN, C.T. Learning to Control Heart Rate: Effects of Varying Incentive and Criterion of Success on Task Performance. Psychophysiology 13, 1976, 378-384
- LANGER, K. Studie zur Problematik optimaler Gruppierung. Dipl.Arbeit an der WWU-Münster 1975
- Zielgruppenanalyse - Zur Auswertung gesellschaftlich-politischer Daten mit Hilfe von Gruppierungsverfahren. Dipl.Arbeit an der WWU-Münster 1975
- LA SALLE, J. & LEFSCHETZ, S. Die Stabilitätstheorie von Ljapunov. Mannheim 1967
- LAURIG, W. & PHILIPP, U. Veränderungen der Pulsfrequenzarrhythmie in Abhängigkeit von der Arbeitsschwere. Arbeitsmed., Sozialmed., Arbeitshyg. 5, 1970, 184-188
- LAWELY, D.N. & MAXWELL, A.E. Factor Analysis as a Statistical Method. London 1971²
- LEGEWIE, H. & EHLERS, W. Knaurs moderne Psychologie. München 1972
- LEGEWIE, H. & NUSSELT, L. (Ed.) Biofeedback-Therapie. München 1975
- LEVITT, E.E. Die Psychologie der Angst. Stuttgart 1976

- LIGHTHILL, M.J. Einführung in die Theorie der Fourier-Analyse und der verallgemeinerten Funktionen. Mannheim 1966
- LIPTSER, R.S. & SHIRYAYEV, A.N. Statistics of Random Processes II: Applications. Berlin 1978
- LUCZAK, H. & LAURIG, W. An Analysis of Heart Rate Variability. Ergonomics 16, 1973, 85-97
- LUCZAK, H. & RASCHKE, F. Regelungstheoretisches Kreislaufmodell zur Interpretation arbeitsphysiologischer und rhythmologischer Einflüsse auf die Momentanherzfrequenz: Arrhythmie. Biol.Cybernetics 18, 1975, 1-13
- MATSUSHITA, Y. Hidden Lines Elimination for a Rotating Object. Comm.ACM 15, 1972, 245-252
- Mc FARLAND, R.A. Heart Rate Perception and Heart Rate Control. Psychophysiology 12, 1975, 402-405
- Mc GRATH, F.J. A Method for Eliminating Hidden Lines with Polyhedra. Simulation 16, 1971, 37-41
- MEISEL, W.S. Computer-Oriented Approaches to Pattern Recognition. New York 1972
- MESCHKOWSKI, H. Wahrscheinlichkeitsrechnung. Mannheim 1968
- MICHAELIS, J. Zur Anwendung der Diskriminanzanalyse für die medizinische Diagnostik. Habil.Schrift der Universität Mainz, Mainz 1972
- MILLER, N.E. & DWORKIN, P.F. Recent Difficulties with Curazied Rats and Significant Problems for Human Research. In: OBRIST, P.A. et al (Eds.) Contemporary Trends in Cardiovascular Psychophysiology. Chicago 1974
- MULDER, G. & MULDER-HAJONIDES VON DER MEULEN, W.R.E.H. Mental Load and the Measurement of Heart Rate Variability. Ergonomics 16, 1973, 69-83
- NEUNDÖRFER, B. EEG-Fibel. Stuttgart 1975
- OBRIST, P.A. The Cardiovascular-Behavioral Interaction - As its Appears Today. Psychophysiology 13, 1976, 95-107
- OPMEER, C.H.J.M. The Information Content of Successive RR-Intervall Times in the ECG. Ergonomics 16, 1973, 105-112
- OPPENHEIM, A.V. & SCHAFER, R.W. Digital Signal Processing. New Jersey 1975

- OTNES, R.K. & ENOCHSON, L. Digital Time Series Analysis. New York 1972
- PAWLIK, K. Dimensionen des Verhaltens. Bern 1968
- Psychologische EEG-Forschung. Methoden, Ergebnisse, Anwendungen. In: SCHRAML, W.J. & BAUMANN, U. (Ed.) Klinische Psychologie II. Bern 1974, 416-456
- PAPOULIS, A. The Fourier Integral and its Applications. New York 1962
- Probability, Random Variables, and Stochastic Processes. New York 1965
- PICKENHAIN, L. Methodologische Probleme der Untersuchung biologischer Faktoren bei psychiatrischen Erkrankungen. In: PICKENHAIN, L. & THOM, A. (Eds.) Beiträge zu einer allgemeinen Theorie der Psychiatrie. Jena 1968
- PONGRATZ, L.J. (Ed.) Handbuch der Psychologie. Bd. 12: Klinische Psychologie, 2. Halbbd. Göttingen 1978
- RABINER, L.R. & GOLD, B. Theory and Application of Digital Signal Processing. New Jersey 1975
- RAND, W.M. Objective Criteria for the Evaluation of Clustering Methods. J. Amer. Stat. Assoc. 66, 1971, 846-850
- RAO, C.R. Recent Trends of Research in Multivariate Analysis. Biometrics 28, 1972, 3-22
- REINSCH, C.H. Smoothing by Spline Functions. Num. Math. 10, 1967, 177-183
- REVENSDORF, D. Faktorenanalyse. Stuttgart 1977
- ROBINSON, J.O. The Computer in Clinical Psychology. In: APTER, M.J. & WESTBY, G. (Eds.) The Computer in Psychology. London 1973, 221-245
- ROEDER, B. Die Konfigurationsfrequenzanalyse (KFA) nach Krauth und Lienert. Z.f. SS 26, 1974, 819-844
- ROHMERT, W. et al Heart Rate Variability and Work-Load Measurement. Ergonomics 16, 1973, 33-44
- ROUCHE, N. et al Stability Theory by Liapunov's Direct Method. New York 1977

- RUBIN, J. Optimal Classification into Groups: An Approach for Solving the Taxonomy Problem. J.Theoretical Biology 15, 1967, 103-144
- SARASON, I.G. & SPIELBERGER, C.D. (Ed.) Stress and Anxiety. New York 1975
- SAYERS, B.McA. Analysis of Heart Rate Variability. Ergonomics 16, 1973, 17-32
- SCHEFF, T.J. Die Rolle des psychisch Kranken und die Dynamik psychischer Störung. Ein Bezugsrahmen für die Forschung. In: KEUPP, H. (Ed.) 1972a, 136-156
- SCHLOSSER, O. Einführung in die sozialwissenschaftliche Zusammenhangsanalyse. Reinbeck 1976
- SCHMIDT, R.F. Integrative Funktion des ZNS. In: SCHMIDT, R.F. & THEWES, G. (Eds.) 1976, 146-177
- SCHMIDT, R.F. & THEWES, G. Einführung in die Physiologie des Menschen. Berlin 1976
- SCHÖNPFLOG, W. Methoden der Aktivierungsforschung. Bern 1969
- SCHWARZ, H. Optimale Regelung linearer Systeme. Mannheim 1976
- SELYE, H. The Stress of Life. New York 1956
- The Concepts of Stress in Experimental Psychology. In: TANNER, J. (Ed.) Stress and Psychiatric Disorder. Oxford 1960
- SINGLETON, R.C. An Algorithm for Computing the Mixed Radix Fast Fourier Transform. IEEE Trans.Audio. Electroacust. AU-17, 1969, 93-103
- SKARABIS, H. Mathematische Grundlagen und praktische Aspekte der Diskrimination und Klassifikation. Würzburg 1970
- SMIRNOW, W.I. Lehrgang der Höheren Mathematik. Teil II. Berlin (VEB) 1964
- SODEUR, W. Empirische Verfahren zur Klassifikation. Stuttgart 1974
- SPÄTH, H. Spline-Algorithmen. München 1973
- Cluster-Analyse-Algorithmen zur Objektklassifizierung und Datenreduktion. München 1975
- Algorithm 30: L₁ Cluster-Analysis. Comp. 16, 1976, 379-387

- SPÄTH, H. Numerische Erfahrungen zu heuristischen Lösungsverfahren beim Varianzkriterium in der Cluster-Analyse. Angew.Infor. 1977, 67-72 (1977a)
- Fallstudien Clusteranalyse. München 1977 (1977b)
- SPIEGEL-RÖSING, I.S. Wissenschaftsentwicklung und Wissenschaftssteuerung. Frankfurt 1973
- SPIELBERGER, C.D. (Ed.) Anxiety. New York 1972
- SPRENG, M. Zusatzmethoden zur Verbesserung der Aussagekraft gemittelter evozierter Potentiale. Z.EEG-EMG 3, 1972, 49-56
- SPRENG, M. & KEIDEL, W.D. Problems of Simple Averaging of Electro-Physiological Recordings and the Use of Additional Methods. Rev.Laryng.Supplement 1971, 722-738
- SROUFE, L.A. Learned Stabilization of Cardiac Rate with Respiration Experimentally Controlled. J. Experimental Psychol. 81, 1969, 391-393
- Effects of Depth and Rate of Breathing on Heart Rate and Heart Rate Variability. Psychophysiology 8, 1971, 648-655
- STEINHAUSEN, D. & LANGER, K. Clusteranalyse. Berlin 1977
- SUPPES, P. & ZINNES, J.C. Basic Measurement Theory. In: LUCE, R.D. et al (Eds.) Handbook of Mathematical Psychology. Vol.I. New York 1963, 2-76
- THEWES, G. Lungenatmung. In: SCHMIDT, R.F. & THEWES, G. (Eds.) 1976, 452-486
- THOMPSON, R.F. Introduction to Physiological Psychology. New York 1975
- TRIEBEL, H. Höhere Analysis. Berlin (VEB) 1972
- ÜBERLA, K. Faktorenanalyse. Berlin 1971²
- UNBEHAUEN, H. & BAUER, B. Parameterschätzung zur Systemidentifikation. München 1974
- UNBEHAUEN, R. Systemtheorie. München 1969
- VAITL, D. Zur Problematik des Biofeedback dargestellt am Beispiel der Herzfrequenzkontrolle. Psych. Rundschau 26, 1975, 191-211 (1975a)

- VAITL,D. Biofeedback-Einsatz in der Behandlung einer Patientin mit Sinustachykardie. In: LEGEWIE, H. & NUSSELT,C. (Eds.) 1975, 205-217 (1975b)
- VAITL,D. & KENKMANN,H.J. Stabilisation der Pulsfrequenz durch visuelle Rückmeldung. Z.Klin.Psych. 1, 1972, 251-271
- VAITL,D. et al Heart-Rate Stabilization Feedback and Concomitant Physiological Changes. Unveröff. Manuskript, Gießen 1978
- VAN DEN DAELE,W. & WEINGART,P. Resistenz und Rezeptivität der Wissenschaft - Zu den Entwicklungsbedingungen neuer Disziplinen durch wissenschaftspolitische Steuerung Z.f.Soz. 4, 1975, 146-164
- VENABLES,P.H. & CHRISTIE,M.J. Research in Psychophysiology. London 1975
- VOGEL,F. Probleme und Verfahren der numerischen Klassifikation. Göttingen 1975
- VOS,J.E. Representation in the Frequency Domain of Nonstationary EEGs. In: DOLCE,G. & KÜNKEL, H. (Eds.) Computerized EEG Analysis. Stuttgart 1975, 41-50
- WALSH,J.L. A Closed Set of Normal Orthogonal Functions. Ann.J.Math. 55, 1923, 5-24
- WATANABE,S. Karhunen-Loeve Expansion and Factor Analysis. Theoretical Remarks and Applications. Trans. 4.Prague Conf. 1965 on Inform.Theory, Prag 1967, 635-660
- WEBER,E. Einführung in der Faktorenanalyse. Stuttgart 1974
- WEEDE,J. & JAGODZINSKI,W. Einführung in die konfirmatorische Faktorenanalyse. Z.f.Soz. 6, 1977, 315-333
- WEINBERG,F. Branch and Bound: Eine Einführung. Berlin 1973
- WELCH,P.D. The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging over Short Modified Periodograms. IEEE Trans.Audio.Electroacus. AU-15, 1967, 70-73
- WERNER,W. Vorlesung über Approximationstheorie. Berlin 1966
- Neuere Entwicklungen auf dem Gebiet der nicht-linearen Splines. Schriftenreihe des Rechenzentrums Münster Nr. 26, Münster 1977 (1977a)

- WERNER, H. Algorithms for Numerical Integration with Regular Splines. Schriftenreihe des Rechenzentrums Münster Nr. 77, Münster 1977 (1977b)
- WERNER, H. & SCHABACK, R. Praktische Mathematik II. Berlin 1972
- WIENER, N. The Interpolation, Extrapolation and Smoothing of Stationary Time Series Analysis. New York 1949
- WILLIAMSON, A. Hidden-Line Plotting Program. Comm.ACM 15, 1972, 100-103
- WINKLER, G. Stochastische Systeme. Analyse und Synthese. Wiesbaden 1977
- WIRTH, D. Dynamische Wechselbeziehungen zwischen Herz-Kreislauf-System und Atmung. In: DRISCHEL, H. (Ed.) Einführung in die Biokybernetik. Berlin 1973, 213-245
- WOMACK, B.F. The Analysis of Respiratory Sinus Arrhythmia Using Spectral Analysis and Digital Filtering. IEEE Trans.Bio-med.Eng. BME-18, 1971, 399-409
- WUNDT, W. Grundzüge der physiologischen Psychologie. Bd. I. Leipzig 1874
- ZUCKERMAN, M. & SPIELBERGER, C.D. (Eds.) Emotion and Anxiety. New York 1976

LEBENS LAUF

Am 20. Februar 1951 wurde ich als Sohn des kaufmännischen Angestellten Hans Langer und seiner Ehefrau Eva, geb. Krause, in Soest (Westf.) geboren.

1957 wurde ich in die Volksschule meines Heimatortes eingeschult. Mit dem Unzug der Familie nach Köln wechselte ich im selben Jahr auf die Volksschule der Stadt Köln, die ich zwei Jahre lang besuchte. 1959 kam ich auf die Volksschule der Stadt Münster. In der Zeit von 1962 - 1970 absolvierte ich das Konrad-Schlaun-Gymnasium, an dem ich 1970 auf dem mathematisch-naturwissenschaftlichen Zweig das Abitur ablegte.

Im Wintersemester 1970/71 immatrikulierte ich mich an der Westfälischen-Wilhelms Universität zu Münster in den Fächern Psychologie, Pädagogik und Soziologie. 1972 bestand ich die Vordiplomprüfung in Pädagogik und 1973 die Vordiplomprüfung in Psychologie. 1975 erhielt ich das Diplom in Psychologie und das Diplom in Pädagogik.

Mit besonderem Interesse habe ich desweiteren Vorlesungen und Seminare in den Fächern Mathematik und Medizin besucht. 1977 legte ich die ärztliche Vorprüfung ab.

Neben der Unterstützung, die mir durch die Graduiertenförderung zuteil geworden ist, sind meine Studien und Arbeiten durch die Einrichtungen und Mittel des Rechenzentrums der Universität Münster gefördert worden, an dem ich seit 1977 insbesondere im Rahmen der Forschungsprojektes "Mathematische Modelle in der Medizin" tätig bin.

